IMPROVING DYNAMIC DECISION MAKING THROUGH RFID: A PARTIALLY OBSERVABLE MARKOV DECISION PROCESS (POMDP) FOR RFID-ENHANCED WAREHOUSE SEARCH OPERATIONS

By

SHARETHRAM HARIHARAN

Bachelor of Engineering in Electronics and

Communication

University of Madras

Chennai, India

2004

Submitted to the Faculty of the
Graduate College of the
Oklahoma State University
in partial fulfillment of
the requirements for
the Degree of
MASTER OF SCIENCE
December 8, 2006

IMPROVING DYNAMIC DECISION MAKING THROUGH RFID: A PARTIALLY OBSERVABLE MARKOV DECISION PROCESS (POMDP) FOR RFID-ENHANCED WAREHOUSE SEARCH OPERATIONS

Thesis Approved:

Dr. Satish Bukkapatnam

Thesis Adviser

Dr. Manjunath Kamath

Dr. Tieming Liu

Dr. Gordon Emslie

Dean of the Graduate College

RESEARCH SUMMARY

Radio Frequency Identification (RFID) tags are increasingly being used to track inventory in the supply chain and also in many other areas where unique identification of items is required. Though the basic technology has been in existence for a long time, its application was confined largely to the military sector. Mandates by retailers such as Wal-Mart, Target and from the US Department of Defense are currently acting as a major thrust to the commercial adoption of RFID.

This thesis focuses on formulating and solving decision making models to aid the search for misplaced items in a warehouse or a storage location, using RFID signals from tagged inventory as beacons. However, the uncertain nature of signal strengths received due to the presence of metals and liquids, makes the search process cumbersome. The dense packaging of items also leads to missed reads and bad read rates. Partially Observable Markov Decision Process (POMDP) is a framework to model problems with uncertainties in actions and observations. In this thesis different scenarios with varying observations of the signal strengths (observation probabilities) are considered using a POMDP framework.

Current literature on RFID models (both analytical and simulation models) assume that the items are not misplaced in the storage locations and that RFID has 100% read-rate accuracy. However, this is not true because only in very specific scenarios (e.g.,

conveyor belts) 100% read rates have been achieved that too with a not-so-dense packing of items.

A Forklift Operator (FLO) in a warehouse does not know the exact location of the tagged misplaced item and is guided by the imperfect variations in the strength of the signal received from the RFID tag (active or passive) in response to the reader's query. The POMDP Warehouse Search (POMDP-WS) model considers five actions (corresponding to four rectilinear movements of the FLO, plus one stay-put action). Each action of the FLO leads the transition to a partially known state, which in turn results in the manifestation of one of the following five observations: signal increase, signal decrease, item found, no signal and no change. Based on the observations the belief state of the POMDP that captures the Bayesian probabilities of the item to be at different states- is updated. We evaluated the effects of signal strengths, discount factor and the initial beliefs on the search performance of the POMDP. The POMDP provides shortest path to locate the tag in the excellent observation scenario where the observation probability is close to one for one of the observations and almost zero for all others. As the observation probabilities decrease, i.e., there is more random imperfections in the nature of signals received (observations), the number of steps to reach the tag increases considerably. The expected reward from a 20-step POMDP with reasonable observation probabilities (varying between 70 and 90%) was 56% higher than that for a no-RFID case. This result implies that a significant (~56%) reduction in search times and efforts for locating a misplaced item are possible using RFID in the tested scenarios.

ACKNOWLEDGMENTS

Directions, not solutions, are the key to a strong research motivation in me. That is exactly what my advisor Dr.Satish Bukkapatnam gave me. My master's thesis has been inspired and motivated by Dr.Bukkapatnam. Without our discussions and his ideas of different scenarios to disambiguate the RFID systems, it would have been extremely difficult to achieve this task in a short period of time. He gave me the freedom to experiment and I would like to take this opportunity to thank him sincerely not only for his directions but also for many of his advice- technical and non-technical.

Dr.Manjunath Kamath, who prefers quality to quantity, be it publications or solutions to problems, is definitely my other inspiration. He has a tremendous amount of patience, respect for both students and colleagues (which is not required for a person of his stature) and when it comes to technicalities he never makes many assumptions. I would definitely strive to be as thorough as him.

I have taken Advanced Production Control course under Dr.Tieming Liu and would also like to take this opportunity and thank him for his support and guidance in this research and hope to interact more with him in the future.

Chetan Yadati, Ananth Krishnamoorthy, Karthik Ayodhiramanujan, Sandeep Srivathsan, Jayjeet Govardhan and members of COMMSENS lab have helped me in various aspects of my master's research and I would also like to wish them luck and thank them for all their help.

I wouldn't be here if it were not for my parents who always wished to see a continuous improvement in me. Talking to my grandparents, my uncle, aunt and my sisters encourages me to strive for higher goals. A very famous person once remarked- "Aim for the peak, and if you reach it, aim for the sky!" and that is what I am trying to do.

Currently, RFID systems are viewed with skepticism in the industry because of the cost associated with their implementation and some technical disadvantages. We believe that our research effort would definitely steer away this ship of skepticism and enable other research to complement what we have done. We would also like to thank the National Science Foundation and the Center for Engineering Logistics and Distribution (CELDi) for the research grant.

GLOSSARY OF SELECTED TERMS

GEN2	EPC Global standard for air interface protocol for second generation of EPC technologies
EPC	Electronic Product Code to identify individual physical objects
RF	Radio Frequency-refers to that portion of the electromagnetic spectrum in which electromagnetic waves can be generated by alternating current fed to an antenna
WORM	Write Once Read Many is a kind of tag in which information can be written only once and thereafter can only be read many times
EMI	Electromagnetic Interference is the interference caused when the radio waves of one device distort the waves of another device

Path loss	The attenuation that the signal undergoes in traveling over a path between two points. It varies inversely as the square of the distance traveled
POMDP	Partially Observable Markov Decision Process
FLO	Forklift Operator

TABLE OF CONTENTS

LIST OF FIGURES

Figure Page

LIST OF TABLES

CHAPTER I

INTRODUCTION AND RESEARCH OBJECTIVES

1.1 Introduction

In many industrial sectors (e.g. automotive, electronic goods, consumer products, etc.) there is an increasing demand for custom built products. Increasing product variety and inventory levels in a warehouse or a retail store leads to increasing number of misplaced items [1]. If the misplaced items are not found within a specified time they lead to lost sales. Also if the items are perishables the problem becomes even more severe.

A significant percentage of lost sales are due to misplaced or "undetected" items present in the warehouse or retail (back room) storage. The issue of misplaced items maybe addressed by initiating a search process to locate them quickly and efficiently without wasting the consumers' time. The speed of the search process is the key to ensure profitability for the store or the warehouse in terms of operating efficiency. Radio Frequency Identification (RFID) could be a solution to this problem of locating the misplaced items in addition to the inventory accuracy problem in the supply chain. RFID signals can potentially serve as beacons to search for misplaced items. However, given its current limitations, the adoption of RFID for this application has been slow. This research aims at addressing these problems and enables a forklift operator (in a warehouse) or the clerk (in a retail store) with a mobile RFID reader, to quickly locate the specified product(s) in a dense storage environment.

Since RFID signals (like any other electromagnetic signal) are subject to degradation due to noise and other external factors, it is very important to model a system which may not be 100% accurate in its prediction of the location of an item. This is because the signal strength of a signal is affected by (among other issues) distance, interference from other frequency sources and presence of materials that inhibit an efficient transmission of the signal from a source to its destination. This can be captured by a model in which the information of the tag's location is not available accurately to enable the forklift operator to make decisions as to which direction to take to reach the tag.

1.2 Research Objectives

In this research we aim to model, evaluate and compare the process of searching for misplaced items in a warehouse or a storage facility using RFID systems with different signal strength observation probabilities. A model that captures the uncertainty of the signal strength observations at different locations is required to compare and evaluate the search process. Partially Observable Markov Decision Process (POMDP) is a framework which enables to model this scenario with incomplete information.

The specific tasks are structured below:

1. Literature survey of warehouse operations and the effect of misplaced items resulting in stock loss and stock-out
2. Literature survey of decision making approaches using RFID information

3. Literature review of MDP and POMDP to identify model parameters
4. Evaluate, through simulation study, the effect of RFID information accuracy on improving inventory accuracy
5. Formulate searching of misplaced items as a POMDP and incorporate the detrimental effects of false negative and false positive reads the model
6. Evaluate the POMDP with RFID information and compare expected rewards and policies for different types (three) of signal strength observations in a warehouse

The remainder of this thesis is organized as follows: Chapter 2 introduces the RFID technology in detail with its advantages, disadvantages and applications. It also explains the need for accurate inventory records and the effects of the lack of it with examples from industry case studies. In the second half of the chapter, we have provided a brief description of the decision-making approach using Markov Decision Processes (MDPs) and the elements involved in describing the MDPs. Finally we conclude with a description of the Partially Observable Markov Decision Process (POMDP), partial observability in RFID and the effects of false negative reads ad positive reads on the inventory records. Chapter 3 contains the current literature of RFID analytical and simulation models, their gaps and the thesis approach. Chapter 4 explains the formulation of the POMDP Warehouse Search (WS) model for different observation probabilities. Chapter 5 explains the results of the different models and the corresponding policies and expected rewards obtained for different belief states. Finally we conclude the thesis in Chapter 6 summarizing the contributions of this thesis and ideas for future work.

CHAPTER II

BACKGROUND

This chapter presents a review of the literature on POMDPs and warehouse operations with RFID. We begin with a basic introduction to the technology of RFID, its advantages and disadvantages in Section 2.1 and then proceed to warehouse management operations and use of RFID in minimizing inventory stock loss and stock outs in Sections 2.2 through 2.5. Then we present a brief primer on Markov decision processes and move on to partial observability in RFID in the Sections 2.7 through 2.10.

2.1 Radio Frequency IDentification

RFID stands for Radio Frequency IDentification, which is a tracking technology aimed at attaining very high real-time visibility of a product in a supply chain. Though the above definition states supply chain as the main application, RFID is now being used, quite successfully, in many areas including inventory control, cattle tracking, pharmaceutical industries, safety devices in factories, hospitals/health-care, tire manufacturing, libraries, airport security/ baggage handling, container/pallet tracking, stolen vehicle identification, car body production, oil pipe marking and toxic waste monitoring.

The current RFID systems consist of tags (active, passive, semi-passive), readers, antennae and computer systems called middleware which connect to the back-end databases of applications (e.g. ERP systems). A reader is capable of generating and receiving RF signals. As shown in Figure 1, the reader sends RF signal into the environment. This electromagnetic signal is a query from the reader to the tag to identify itself and provide other information about the product it is on. As the tag comes into the reader's electric field, the tag circuit sends signals back to the reader thus identifying the object. The tags and the reader communicate multiple times to ensure secrecy of information and hence also prevent error in transmission. This technology can be used for real time job tracking, goods and asset management, etc. Currently tags maybe active, passive or semi-passive. They are typically read-only, read-write, write-once read-many (WORM) or write-many read-many. The readers too have different specifications like frequency, type of data transmission method, etc., which determine the performance of a system.

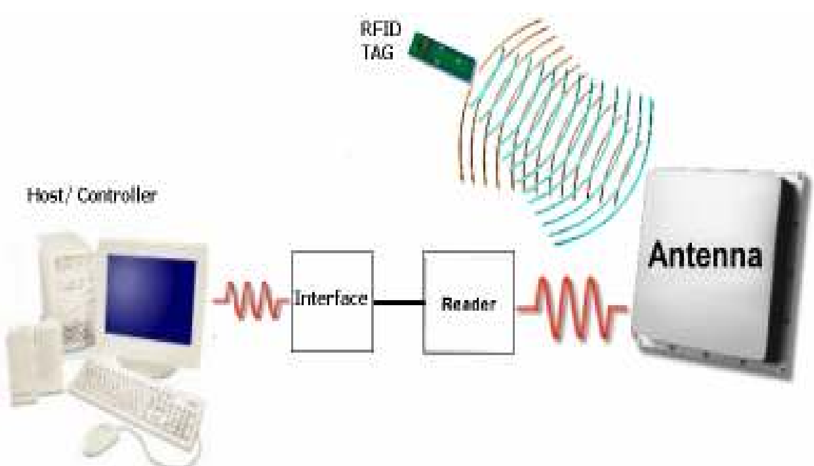

Figure 1: A simple RFID system

2.1.1 Advantages of RFID

Some benefits due to the adoption of automatic identification technology are described below [2]:

1. Reduction in labor costs by minimizing the number of operators needed for scanning at each stage in the supply chain
2. Reduction in queue length and customer waiting time due to acceleration of physical flows- this is due to the large number of items scanned simultaneously by an RFID reader
3. More efficient control of supply chain operations due to increased information accuracy
4. A thorough knowledge of out-of-stock situations due to continuous monitoring of items- Procter and Gamble reported that out-of-stocks in retail stores caused about $ 3 billion in lost revenue for a year and on 10% of the shopping trips the consumer cannot find what he/she wanted and bought something else or nothing at all [3]
5. Increased management of returned items, perishable goods, quality assurance processes among others is ensured with RFID adoption

2.1.2 Current disadvantages of RFID

RFID has its own disadvantages amidst the myriad of benefits that it seems to provide. Some of the *current* disadvantages of RFID are:

- The cost of an RFID tag is high – it must be as low as $0.05 for justifying an ROI (Return On Investment) according to RFIDJournal.com

- Effect of metals and fluids- Radio Frequency waves are reflected by metals and absorbed by fluids and RFID tends to perform poorly in the presence of these
- Lack of standards for global adoption
- Read range: The passive tag can be read over a very short distance, typically up to 8 and, in ideal case, at most 12 feet
- Susceptibility to electromagnetic interference (EMI)
- Information overload: The amount of information that gets generated implementing RFID and hence data management becomes a more critical issue

2.2 Inventory and Inventory Records

The amount of raw materials, work in process, and finished goods being held for sale at a given time for a company is known as inventory in warehouse literature. Inventory records are hard copies or electronic documents that reflect how much and what kind of inventories a company has on hand, committed (allocated) to work-in-process, and, on order.

2.2.1 The need for accurate Inventory Records

Some of the reasons cited for the need of accurate inventory records in [4] are as follows:

1 For financial planning including cash-flow analysis, year-end tax calculations, and financial reports

2 For Marketing and Sales planning to sell existing inventory

3 For planning the launch of new products

4 For procurement planning using Material Requirements Planning (MRP)

5 To satisfy the U.S. government's Material Management and Accounting Systems Standards (MMASS) which state that 95% inventory record accuracy is desirable for contractors and subcontractors (Standard #5)

2.3 Stock loss and Stock out

Kang and Gershwin [5] give a comprehensive analytical/simulation study that describe the effect of stock loss on inventory accuracy and compare the various approaches including the use of Auto-ID technologies to resolve the stock loss and stock out problem. In essence, even a small rate of stock loss that is undetected by the information system or the Inventory Manager (IM) could result in severe out-of-stocks and disrupts the entire replenishment process. Retailers, according to their studies, are not aware of the number of products in their stores. Further, they conducted a study of a company which carries thousands of products which translates into a large number of SKUs. On an average, the inventory accuracy of the company stores is about 51%. This number was based on the parameter called *perfect inventory accuracy* which is defined as the percentage of SKUs whose inventory record matches actual stock *perfectly*. Even by relaxing the constraint of perfect match to ± 5 deviations in number the average accuracy for that company is only about 76%.

A similar finding has been reported in California Management Review [6]. The investigation of 370,000 SKUs, about 65% of the inventory records, did not match the physical inventory at the store-SKU level. In addition, 20% of the inventory records differed from the physical stock by greater than or equal to 6 items. The company had an automatic replenishment process with the help of information technology in this case.

2.3.1 Causes of Inventory Inaccuracy

Kang and Gershwin's paper also elicits some of the commonly observed causes for discrepancies in the records: stock loss, transaction error, inaccessible or misplaced inventory and incorrect product identification.

Causes of stock loss include the following: theft (both by employees and shoppers), out-of-date or spoiled/damaged products, and *misplaced items*. Shipments that arrive from suppliers and checkout shipments from warehouses/DCs and stores have transaction errors in the form of discrepancy between shipment record and actual shipment. When products are present in the storage aisles either in the stores or in the warehouses but are not available for sale because they are misplaced, it is equivalent to a *lost sale*. Additionally if the item is perishable or has a short expiry date then they could be branded as a lost sale again. Loss of goodwill due to non-availability of a product when it is actually misplaced is extremely difficult to quantify but definitely has a great impact on the business.

Another significant case study reported in their paper is that of the ECR Europe [7]. Stock loss for 200 companies of the consumer goods category amounts to 1.75% of annual sales for the retailers. This is equivalent to a whooping 13.4 billion euros annually. Of this 59% was unknown to the retailers, meaning, the retailers were clueless as to where or how the products were lost.

2.3.2 Effect of misplaced items and inaccurate information

In a Harvard Business School study, Ton and Raman [1] showed that increasing product variety and inventory level per product results in an increase in misplaced products. The consequence of this is a decrease in the sales of the retail store. 333 stores of the chain Borders Group Inc. were used to establish their claims on product variety and misplaced items. The study was conducted over a period of 4 years and was motivated by “phantom stockouts”. Phantom stockouts are situations in which customers are unable to find products that are actually present (available) in the stores. Based on an exploratory study they also reported that one out of six customers did not find the item because they were not “out of stock” but were placed in backrooms or other storage locations.

Andersen Consulting published a report in 1996 which estimated that sales lost due to products that were present in storage areas but not on the selling floor amounted to $ 560-960 million per year in the US supermarket industry. Amazon.com had about 12% of its inventory stored in the wrong places in 2000 which it then claimed to have brought it down to 4% in 2002 according to a Business Week (2002) report. For an online retailer like Amazon, 4% is still a huge number considering the volume of inventory handled.

The results published by Ton and Raman are two fold. Increasing product variety and inventory level per product at a store leads to an increase in misplaced products. Misplaced products lead to lost sales which affect store profitability. Also, their tests provide empirical evidence to support assertions that higher product variety and inventory levels lead to an increase in defect rates and that increased quality benefits a firm’s performance financially.

2.4 A brief primer on warehouses and operations

Warehouses play very important roles in the supply chain. Some of the common types of warehouses are raw materials warehouse, finished goods warehouse and work-in-process warehouse. In addition some warehouses also act as distribution centers or fulfillment centers in some supply chains.

Though there are different types of warehouses, most of them have the following fundamental set of common operations, though the operational methodology and technology used may vary from one to another:

- Receiving
- Prepackaging (optional)
- Putaway
- Storage
- Order picking
- Packaging and/or pricing (optional)
- Sorting and/or accumulation
- Unitizing and shipping

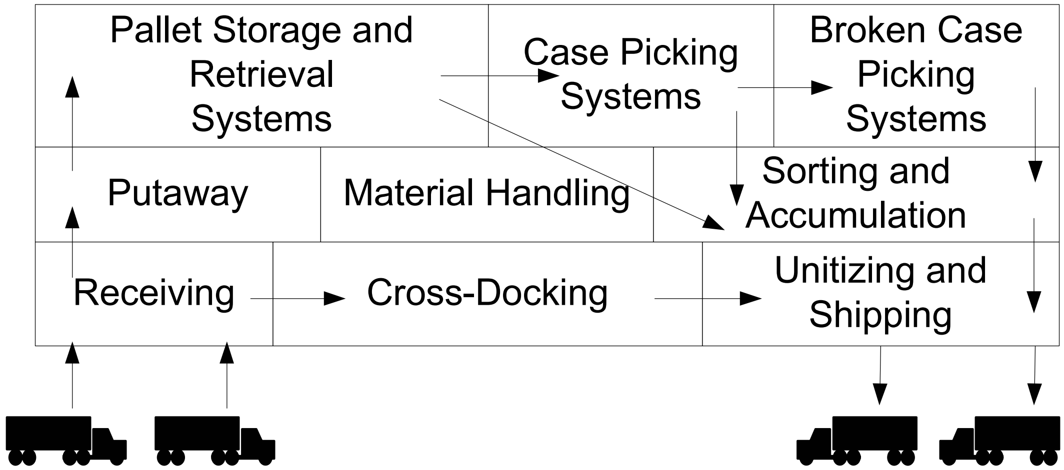

Figure 2: Common warehouse activities [8]

Zeimpekis at al. [9] suggested the following warehouse scenario for day to day transactions/operations processing as summarized in Figure 2.

- Receiving docks receive the cartons or pallets where they are unloaded
- Quantities are verified by the warehouse operators by using their bills-of-ladings or manifests
- At the same time random quality checks are performed on the delivered loads
- The loads are quickly calculated for the staff to determine the number of pallets needed for transporting the goods to the storage area
- The goods are then palletized and then a label is generated and attached to each load indicating its assigned location
- Reassembling of the entire incoming stuff is done to adjust to the internal operations of the warehouse
- The goods are again transported to a location (staging) within the storage area

- Operators get to manually enter the temporary holding location within the storage area
- Order Picking is carried out whenever a new order is received from a customer. The warehouse operator must key in details to the central database system to find out the location and the availability of the item
- An order contains the products and quantities requested by a customer or a production/assembly station in the case of a distribution center or a production warehouse
- SKUs or Stock Keeping Units – SKUs refer to a specific item in a specific unit of measure. For example, if you distributed thirty-weight motor oil in both quarts and gallons you would maintain the inventory as two SKUs even though they are both thirty-weight motor oil. Also refers to the identification number assigned to each SKU
- When an order has multiple SKUs these must be sorted and accumulated before being transported to the shipping area or to the production floor
- Accumulation and sorting could be performed after the order picking process or before the process
- Finally the products are shipped from the shipping area after being retrieved from the storage area

2.5 Current practices and limitations of inventory management in warehouses

Given the uncertainties like order cancellation, order modification, expediting orders; supply chains must be modeled as flexible as possible to accommodate them. Today's practices in warehouse environments include manual paper based order verifications or wireless scanners which are hand-held. Companies are on a constant look out for better technology which tracks real time information and uses that information to make meaningful decisions like rerouting certain supplies, canceling orders in real time, ordering instantly to the manufacturer itself etc. This is also called as the 'need for real-time inventory management systems'.

Although the process described above has shades of delivering value to the customer/enterprise and that it is faultless and streamlined, there are many ineffective characteristics when it comes to flexibility. Some of the applications of RFID could be used in:

A. Ability (in real time) to verify the quantity of received goods
B. Label generation i.e. associating the arrival of a product with its assigned location
C. "blind periods" in which the location of an object is unavailable until it is found
D. Misallocation of pallets by the truck drivers due to improper information availability i.e. real time verification of the item and its location
E. Lack of real time connectivity between warehouse and customer

2.6 Markov Decision Processes (MDPs)

Markov Decision Processes are models for sequential decision making when outcomes are uncertain. They are also called stochastic dynamic programs or stochastic control problems [10]. An MDP model consists of 5 elements:

- Decision Epochs
- States
- Actions
- Transition Probabilities
- Rewards

A decision maker, agent or controller can influence the behavior of a probabilistic system as time evolves. He does this by making decisions or choosing actions. He is given the task of choosing the sequence of actions by which the system must perform optimally with respect to some predetermined performance criterion. The decision that we are going to take tomorrow (future) is influenced only by today's (current) decision. Decisions cannot be made at random but must be made to calculate or take into account the future costs (or rewards).

2.6.1 Decision Epochs and Periods

Decisions are made at points of time called decision epochs. T is denoted as the set of decision epochs and it could be continuous or discrete. If T is an interval we denote it by $T= [0, N]$, $N <= \infty$. If N is finite then the problem is finite horizon problem else it is an infinite horizon problem. The last decision is made at decision epoch $N-1$ [10].

2.6.2 State and Action sets

At each decision epoch the system occupies a state and we denote the set of all possible system states by S. If, at some decision epoch, the decision maker observes the system in state $s \in S$, he can take an action a from a set of allowable actions in state s, A_s. We assume that A_s and S does not vary with T which means that S and A are finite and discrete (finite or countably infinite). Actions could be chosen either randomly or deterministically. Choosing actions randomly means selecting a probability distribution [10].

2.6.3 Rewards and Transition Probabilities

As a result of choosing action a in state s at decision epoch t, the decision maker receives a reward $r(s, a)$. The system state j is reached at the next decision epoch with an associated probability given by the transition probability distribution $P_t(j| s, a)$. When the reward depends on the state of the system at the next decision epoch then we denote it by $r_t(s, a, j)$ where j is the next state at epoch $t+1$.

The expected value of the reward at decision epoch t may be computed by the following formula

$$R_t(s, a) = \sum_{j \in s} r_t(s, a, j) P_t(j \mid s, a)$$

In finite horizon MDP no decision is made at decision epoch *'N'*. Therefore the reward at this time point is only a function of the state. It is denoted by - $r_N(s)$ also called the *salvage value or the scrap value* [10].

2.6.4 Representation of an MDP

A Markov Decision Process is described as a 4-tuple $< S, A, T, R >$, where

- S is a finite set of the states of the world

- *A* is a finite set of actions
- *T: S X A → Π (S)* is the state transition function, giving for each world state and agent action, a probability distribution over world states (we denote *T (s, a, s')* for the probability of ending in state *s'*, given that we start in state *s* and take action *a*)
- *R: S X A → R* is the reward function, which gives the expected immediate reward gained, for taking each action in each state. It is universally denoted as *R (s, a)* where taking action *a* in state *s* fetches us the reward (or cost)

2.7 Value Iteration (VI) algorithm for finite state-space MDPs

To compute an optimal policy for MDPs there are various algorithms but since POMDPs also use Value Iteration (VI) to find an optimal policy, an introduction to the algorithm is provided. The optimal policy is achieved by computing a sequence V_t of finite horizon optimal value functions. A VI algorithm (refer Figure 3) makes use of an auxiliary function $Q_t^a(s)$ which is the t step value of starting in state *s*, taking action *a*, then continuing with the t -1 step non-stationary policy . γ is the discount factor which varies between 0 and 1. It is used to ensure that the reward in the infinite lifetime is a finite sum.

The algorithm terminates when the maximum difference between two successive value functions is less than ε, called the Bellman error magnitude.

$V_1(s)$: =0 for all s

t: =1

repeat

t: = t +1

loop for all s Є S

loop for all a Є A

$Q_t^a (s): = R(s, a) + \gamma \Sigma_{s' \in S} \ T (s, a, s') V_{t-1} (s')$

end loop

$V_t(s): = \max_a Q_t^a (s)$

end loop

until $| V_t(s) – V_{t-1} (s') | < \varepsilon$ for all s Є S

Figure 3: Value Iteration algorithm for MDP

2.8 Partial Observability and motivation for POMDP

For an MDP *each state must be completely observable i.e. its value must be known at all times*. But in a POMDP the knowledge of a current state is non deterministic. Since the current state itself is uncertain, *actions based on the current states are no longer valid*.

The main difference between the MDP and a POMDP is the concept of 'observability'. The state gives us an observation that can provide a hint as to what state the system is currently in. Since this observation can be probabilistic, we also add *'observation functions'* (OF) that tells us the probability of each observation for each state in the model [11].

2.8.1 Representation of a POMDP

A Partially Observable Markov Decision Process can be described as a 6-tuple [11] $< S, A, T, R, \Omega, O >$, where

- S is a finite set of the states of the world
- A is a finite set of actions
- $T: S \times A \rightarrow \Pi(S)$ is the state transition function, giving for each world state and agent action, a probability distribution over world states (we denote $T(s, a, s')$ for the probability of ending in state s', given that we start in state s and take action a)
- $R: S \times A \rightarrow R$ is the reward function, which gives the expected immediate reward gained, for taking each action in each state. It is universally denoted as $R(s, a)$ where taking action a in state s fetches us the reward (or cost)
- Ω is a set of observations the agent can experience of the world it is in; and
- $O : S \times A \rightarrow \Pi(\Omega)$ is the observation function which gives a probability distribution over possible observations, for each action and resulting state (it is generally denoted by $O(s', a, o)$ for the probability of making observation o given that the agent took action a and landed in state s')

A POMDP is an MDP in which the agent is unable to observe the current state but makes an observation based on the action and the resulting state [11]. An external agent (robot or a human being or in the present context the forklift operator that uses a truck fitted with an RFID reader) makes observations and generates actions. The agent keeps an internal belief state *b,* which summarizes its previous experience in the form of a probability distribution function on different states based on the previous state, action and the resulting observation. The belief state is updated at every epoch using the Bayesian probability equations using the last action, the current observation and the previous belief state. In POMDP a policy is a function of the belief state rather than the actual state of the world. Hence, a belief state maybe viewed as a probability distribution over the states of the world. Also, [11] also claim that the belief state is a sufficient statistic for the past history and initial belief state of the agent. Assuming that the current belief state is computed correctly, we do not need to have any other information such as past actions or observations, to identify or know the current state of the world. This also makes the process over the belief states Markovian.

2.9 Partial observability in RFID

The idea of partial observability in a particular RFID enabled scenario has been discussed by Brusey et al. [12]. They discuss a robotic storage stack and a medicine cabinet (also referred to as the smart cabinet) fitted with RFID readers and tagged parts. In general there are two types of undesirable effects from readers and tags- false negative and false positive reads:

1 **False Negative reads** – the tag might not be read by the reader (even though it may be well within the read range of the scanning reader) and hence we believe that there is no tag in the read range. RF collisions, metal shielding and RF interference from other sources are some of the possible causes

2 **False Positive reads (phantom reads)** - tags might be read when they are outside the scan-region (normally associated with the location of the RFID reader), which makes us believe that the object is present when it has actually left the location to another.

Brusey et al. [12] also propose the use of time based filters to deal with the above scenarios. Consider a real world system with finite number of states at any point in time. They call them the "source" of the signal as some information about its state is received. RF or other sensors receive the information (signal) from such sources in each state. *A fully observable system* is one in which we are able to identify the exact state of the system without any conflicts from the signal itself and *Partial observability*, is when the exact state of the system is unknown from the sensor signal. Most of the real world applications are partially observable in nature.

Two case studies depicting the uncertainty of the reads are described below –one for false positive and another one for false negative. Although false positive reads are less common in real world nevertheless it exists where the metallic environments are close to the readers. Examples could include meat processing units, storage bins or racks that are closed on five sides with metal sheets or trucks that are parked close to the read zone.

2.9.1 False Negative reads (FNR)

Case 1- (Smart Medicine Cabinet) [12]: It is called 'Smart' because of its ability to identify its contents and those that are added or removed. Due to closely placed tags in the medicine cabinet false negative reads occur due to three reasons. First reason is the occurrence of RF collisions or electromagnetic interference because of which some tags are not detected even if they are present in the read range of the reader. The second is due to RF interference or metal shielding and the third reason is due to problems with the RFID reader itself which prevents it from reading once and doing it with 100% reliability. Typical read rates vary depending on the type of use it is put into. RFID readers used to read objects on conveyors might have a 100% read rate because the items are streamlined and not too close to each other. But if they are placed in a warehouse the read rates may drop down to as low as 80% depending on the type of product they are tagged on. This cabinet was attached to a voice recognition system which detected the type of medicine that was placed or removed.

2.9.2 False Positive reads (FPR)

Case 2- (Robotic storage rack) [12]: An experiment was conducted at Auto-ID labs in MIT to demonstrate the integration of RFID based control systems in an automated manufacturing environment. A robot packed men's accessories such as foam, razor, gel and deodorant into a container. Items are pulled from the bottom and packed into empty gift boxes (they are stacked one above the other in 4 different arrays) and the robot packed according to the EPC that it detected. The key operation was sorting the contents according to the EPC and placing the correct product type in the appropriate container. The problem was that the reader read the items above the bottom of the stack which was due to interference from other readers and also due to the shape of the item itself which offsets the RF fields in some cases. The biggest disadvantage due to false positive reads is that if items above are read then the system assumes that the bottom one has been removed which is incorrect.

To solve this problem the authors suggested different methods including using a time based filter with a top hat function which filters events that are Δt_{hat} seconds older than the current time t_{now}. But they acknowledged the fact that there was no theoretical basis for designing the filter and that the filters may not be suitable in other situations. Also the filters had to be tuned which could only be done by trial and error. They also suggest that by modeling the problem as a Partially Observable Markov Decision Process (POMDP) one could incorporate the uncertainty involved in the scenario and obtain the solution for it.

2.10 Possible effects of FNR and/or FPR on Inventory Records

- One obvious problem is the exact count of the number of items present in the read range of an RFID reader. The number of items to be entered into the WMS software plays a critical role from information sources like a storage stack in a warehouse or a conveyor in a distribution center (DC).
- If the items are high-value low volume category then the exact count of the item in the entire supply chain is even more important. Deviations in the number could mean huge losses to the company as was the case for Procter and Gamble [3].
- Exact count is more important where company-specific reorder policies are followed. The shift in the reorder point due to inventory record inaccuracy and its effect on stock outs and stock loss has been discussed in detail by Raman [1].

2.11 POMDP in a storage environment

A warehouse is a complex structure into which a partially observable model could fit in well. There will be places (aisles or locations) where there is a constant deterministic observation based on the tag's location. This could be in the conveyors or in the receiving and shipping docks where each unit of the item is scanned as it moves past the readers. However, to search for an item which is misplaced in the temporary or permanent storage areas poses a great challenge due to the following reasons:

1. A lower density of items stored might result in signals being detected even outside the read range of the reader (False Positives). This could be due to the fact that there might be unfilled spaces between the racks that make the signal traverse a larger distance.
2. A higher density of items stored might result in signals being scattered (or absorbed) and not received by the reader (False Negatives) even if the tag is well-within the read range of the reader
3. The packaging effects, as outlined earlier, play a crucial role in detection/scattering of signals. Materials such as plastics, metals (or a combination of these) and items with liquid content might lead to different signal strengths of the received signals at the same place at different instances of time
4. The racks themselves are metals (in most cases) and thus lead to increased scattering of signals

CHAPTER III

RESEARCH GAPS, PROBLEM DESCRIPTION AND APPROACH

This chapter is divided into three sections of which Section 3.1 deals with current literature on RFID analytical and simulation models. Sections 3.2 and 3.3 illustrate the gaps in these models and the thesis framework, which aims at addressing these gaps, respectively.

3.1 RFID analytical and simulation models

In this chapter we review current research specifically based on analytical and simulation models involving RFID technology.

Lee, Cheng and Leung [13] studied the impact of RFID on supply chain dynamics. Their simulation model provided a quantitative analysis to demonstrate the potential benefits of RFID in inventory reduction and service level improvement. They also argued that the benefits brought about by process transformations enabled by RFID were not truly captured by traditional ROI-type analysis.

The simulation model was developed for a three echelon supply chain model consisting of a manufacturer-DC-retailer setting. They analyzed the effect of three major factors that RFID impacts most-

a) Inventory accuracy

b) Shelf replenishment policy

c) Inventory visibility throughout the supply chain

They have assumed that RFID information is 100% accurate and hence the physical inventory is same as that of the system inventory.

Three case studies with each one addressing each of the above factors were simulated. In each case study, there were two major comparisons- one was a system without RFID and the other was with RFID. There were also cases in which the with-RFID case has different sub-categories such as changing (s, S) policies and store shelf replenishment policies. In addition, the third factor involved different manufacturing quantities, target inventory levels and backorder levels in the DC. Their simulation studies [13] demonstrated that RFID technology has opportunities to provide significant benefits in the supply chain.

Lee and Ozer [14] debate that most of the industry white papers and reports on the value of RFID are not model-based analyses but just educated guesses. Hence, there exists a huge credibility gap of the true value of RFID. They give various examples where there are conflicting values for RFID-supply chain metrics (examples would include reduction in cycle counting costs, labor savings, stocking, distribution, forecast error, inventory discrepancy-values that deviate by a huge margin) in each of the white papers and industry reports they reviewed.

They make two important observations before proceeding with their model. Retail environments supposedly accumulate much more inventory discrepancy (the difference between actual on-hand inventory and the inventory record reported by inventory management systems) because they have high inventory turnovers and more contact with customers than the distribution centers (DCs). The recent developments in information technology have not yet eliminated the inventory discrepancy problem.

The authors claim that to analyze the true value of RFID, factors such as shrinkage, misplacement and transaction errors need to be considered jointly and not in isolation. According to them, these three factors constitute a demand stream that result in inventory discrepancy. The authors explicitly model and incorporate the three demand sources for discrepancy in a finite-horizon, single-item, and periodic-review inventory problem.

Chow et al. [15] developed a design for an RFID case-based resource management system for warehouse operations. An integer linear programming model using a branch and bound algorithm to define the optimum travel distance for forklifts was developed. They have concentrated on a case retrieval and matching process incorporating RFID technology information.

To calculate the picking sequence and travel distance for each of the material handling equipment models, an LP model using the branch and bound method is used. The goal of the objective function is to *optimize the resource model's picking sequence and travel distance* according to the authors. To validate this RFID-RMS (RFID-Resource Management System) a company named GENCO piloted a similar system in one of its distribution centers. The RFID readers were fitted in the forklift trucks and the passive RFID tags were stuck on sides of dock doors and on pallets. These tags gathered data on location of the pallets. The Forklift Operator (FLO)'s location was immediately transmitted to the database with the help of active ultra wideband tags fitted to them. To provide full RF coverage UWB (ultra wideband) tag receivers were placed at strategic locations throughout the warehouse.

Using the nearest neighbor algorithm and inductive indexing approach the case retrieval process is performed by the resource management software. This is done especially to retrieve similar cases. Then these cases are ranked according to 'similarity value' as suggested by the nearest neighbor algorithm. After the manager accepts the ranking order, the next big task was to select the available equipment to perform the order picking process. The real time coordinates of the material handling equipment are transmitted using the UWB antennas. Based on this information the resource management software determines which material handling equipment suits the task best.

The pilot project verified that RFID-RMS enhanced the warehouse operation in GENCO under four major categories of operation level enhancement, operating cost reduction, customer satisfaction and resource management improvement.

3.2 Gaps in models

The current literature of RFID analytical and simulation models do not address the following issues:

1. **Misplaced items:** The RFID-RMS model and the supply chain dynamics model mentioned above assume that the items are not misplaced and hence do not account for the loss due to misplaced items mentioned in the background section. Specifically the RFID-RMS model assumes that items are always available in the racks assigned to them and hence the FLO has to just follow the output of the algorithm in optimizing the travel distance.

2. **Inventory Discrepancy:** As an effect of the false negative reads or system unreliability (or a combination of both), the number of items scanned by the RFID reader could vary under different conditions and hence the exact number of items present in retail storage locations might not be represented accurately.

3.3 Research Approach

3.3.1 Searching for misplaced items in warehouse/retail storage

The problem to locate a particular case in a warehouse with "not so reliable" information from the WMS results in wastage of time, labor and resources. Anecdotal evidence with a government aviation agency (who requested that their name not be cited for security purposes) warehouse operation suggests that the typical search process can take a few weeks to months. If the product is still not found *on time*, one or more of the following may result:

1 lost sales

2 fines for late shipment

3 drop in customer service level

4 indirect costs accrued due to loss of goodwill

5 increased complexity in assembling operations

6 incomplete operations elsewhere

The forklift operator (FLO) enters in his display panel an EPC of the case or pallet that he is searching for and scans the WMS allocated possible location for it. If he/she does not receive any signal (i.e. the product is not present) then he/she moves to another location based on some knowledge. He repeats the process of scanning in any of the four possible directions. If he receives a signal whose strength is limited (there are various thresholds that may classify the signal strength (alternatively, read-rates) as strong, weak, intermediate and so on) then he/she moves in a direction that gives him a better signal strength compared to the other directions. These received signal strengths are called observations. The operator then moves in a direction which translates into updating the belief states in the POMDP domain.

This problem is classified as partially observable because the joint location of the reader *with respect to the tag* (EPC) is unknown. Figure 4 depicts an actual warehouse aisle configuration with a few FLOs:

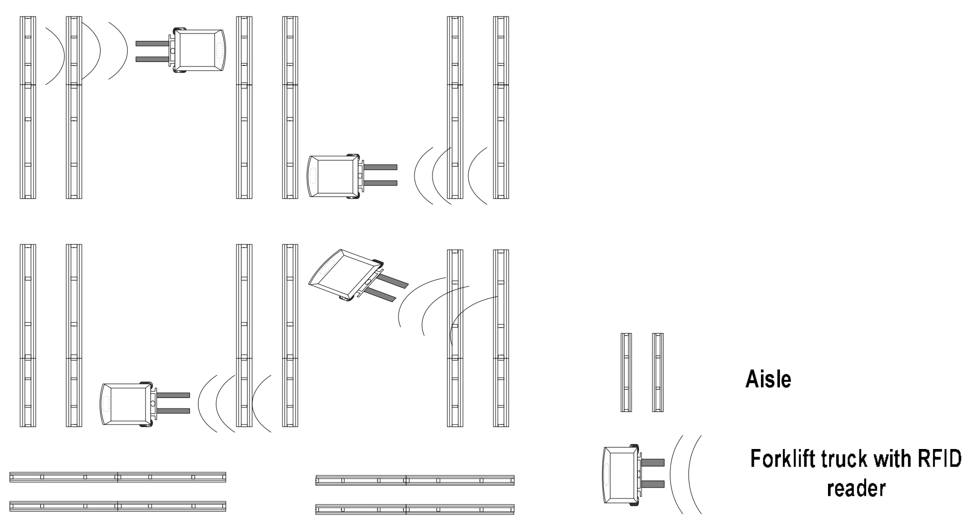

Figure 4: Forklift attached with RFID readers scan aisles

3.3.2 Research Scope and Limitations

This research is focused on trying to find an optimal solution to the misplaced item location problem. However there are certain limitations that need to be specified at this stage. In the Warehouse Search (WS) model formulation in Chapter 4 we model only a small portion of the warehouse in the grid structure with 6 locations. This constraint is placed to overcome the limitations of the POMDP incremental pruning algorithm [16]. If the number of states increases the policy computation time with incremental pruning algorithm becomes prohibitively high. Furthermore, the formulation assumes that only one item needs to be searched at any given point in time and no new search request occurs during the search epoch.

CHAPTER IV

WAREHOUSE SEARCH (WS) MODEL FORMULATION

In many real-world applications the 'state' is rarely completely observable. In addition, we might also not know the effect of an action or decision i.e. the effects of actions may be non-deterministic. In a typical warehouse environment during the order-picking process, a forklift operator might not know the exact location of an item if it was misplaced. Based on the information from a Warehouse Management System (WMS) he/she first reaches a location and scans for its presence. If the item has been moved away from the read range of the RFID reader then he/she does not get any response (signal). But if the item has been moved or misplaced within the read range of the RFID reader, he/she receives a signal based on how far the item is located. The read range of the RFID reader depends on the type of the tag (among other factors) used in the warehouse environment [17]:

- Passive tags- Low Frequency tags have a read range of 0.33 meters (1 foot)
- High Frequency tags have a read range of 1 meter (3 feet)
- Ultra-High Frequency tags have a read range of 10-20 feet
- Active tags have a read range of 100 meters (300 feet)

4.1 Distance and Signal Strength

In telecommunication engineering, the relation between power transmitted from one antenna to another is given by the Friis transmission equation. RFID readers have multiple antennae to transmit power and the tag also has an in-built antenna to process the signal and respond back to the reader. Assuming the distance between the reader and the tag is 'R', we have the following equation

$$\frac{P_r}{P_t} = \frac{G_t G_r \lambda^2}{(4\pi R)^2}$$

where P_r is the power received by the receiving antenna (tag's antenna in our case), P_t is the power input to the transmitting antenna, G_t and G_r are antenna gain of the transmitting and receiving antennas, respectively, and λ is the wavelength [18].

It is quite clear from the above equation that the power received is inversely proportional to the square of the distance between the reader and the tag.

4.2 State space and Belief Space

In the WS model, a 'state' is the joint location of the forklift operator and the tag that he/she is searching for. In a POMDP model a set of states *S = {s1, s2 ...}* describes the problem domain. Referring to Figure 5, if the forklift operator is in location 2 and the tag is in location 3, the WS model system state could be represented as a 2-tuple *(2,3)*. Similarly *(5, 2)* refers to the forklift operator being at location 5 and the tag at location 2.

Hence starting from *(0, 0)* through *(6, 6)* there are 49 states in this WS model as in

$$S = \{(0,0), (0,1), (0,2), \ldots,(0,6),(1,0), (1,1)\ldots(1,6)\ldots(6,6)\}$$

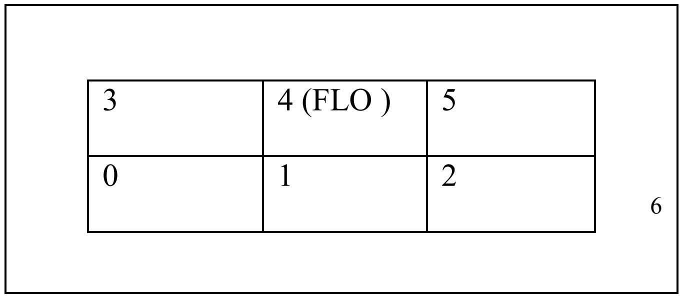

Figure 5: Warehouse Model with wall state

However, a belief state *b* is a probability distribution over the state space *S*. If we denote the probability assigned to world state *s* by belief state *b* as *b(s)*, then

$$0 \leq b(s) \leq 1, \forall (s \in S)$$

$$\sum_{s \in S} b(s) = 1$$

In the WS model if the forklift operator believes that (from the observations in the WS model)the tag is located far away from his current location, i.e., *b* and takes an action *a*, he reaches a next state *b'*. Though the action is deterministic (i.e. if he moves north then he moves north 100 percent of the time) the state that he reaches next is still unclear to him. He is unsure of the tag's location. His/her belief in the new current state *b'* which is *the joint location of the tag and the forklift*, has to be updated. This is done using Bayesian probability theory to update the equation as follows:

$$b'(s') = \Pr(s' \mid o, a, b) \quad = \frac{O(s', a, o) \sum_{s \in S} T(s, a, s') b(s)}{\Pr(o \mid a, b)}$$

This equation defines the update rule for computing a posterior belief, *b'*, given the belief at the previous time step, *b*, and the latest action/observation pair *(a, o).* The denominator is a normalizing factor that makes the *b'* sum to 1. Thus this is the new belief state *b'*. Referring to the WS model, if the forklift operator believes that that tag is in location 3 with probability of 0.6 and that it could be in 0 with 0.4 probability and he/she itself is at position 2, then the current belief state could be represented by:

$$b(s = (2,3)) = 0.6, b(s = (2,0)) = 0.4$$
$$b(s) = 0 \forall s \notin \{(2,3), (2,0)\}$$

4.3 Action/Transition Probability Model

A set of actions *A* describes the possible set of actions the forklift operator can take. In the WS model, we have 5 actions, namely, up, down, left, right and stayput. The decision to choose a particular action at a particular belief state is based upon the action's expected discounted reward starting from that belief state *b* and the observation received from the previous action. Hence the action set *A* could be described as

A = {up, down, left, right, stayput}

4.3.1 State transition probability distribution

If we consider the forklift to be in state *s* and he/she selects an action *a*, the probability that he/she would reach a state *s'* is given by the notation

$$T(s, a, s') = \Pr(s_t = s' \mid s_{t-1} = s, a_{t-1} = a)$$

The implicit assumption here is that

$$\sum_{s' \in S} T(s, a, s') = 1, \forall (s, a)$$

Here, *T(s, a, s')* is called the transition probability function or state-transition function which tells us, for each world state and agent action, a probability distribution over world states. In the WS model the operator knows his/her exact location but is uncertain of the tag's location. But the transition probability matrix captures the movement from one system state to another for each action. In the present case the system state transition function is deterministic. It means that if the forklift is at location 2 and the tag is at 3 i.e. (2, 3) and the action suggested is *up* then the next system state is given by (5, 3). The forklift has moved from location 2 to location 5 and the tag's location remains the same. There is nothing about the movement action of the operator that will cause the tag to move in a probabilistic way.

4.4 Observation Models

An observation model captures the probability that an agent will receive an observation *o*, given that it is in state *s* and has taken action *a*. It is denoted by

$$O(s, a, o) = \Pr(O_t = o \mid s_{t-1} = s, a_{t-1} = a)$$

Similar to the transition-probability function here too we have the implicit assumption that

$$\sum_{s' \in S} O(s, a, o) = 1, \forall (s, a)$$

At any given point in time the system is assumed to be in state s_t which is not completely observable, but is partially observable through observation o_t.

The forklift operator enters in his display panel a tag id (e.g., an Electronic Product Code EPC) of the case or pallet that he is searching for and the WMS displays the possible location. The operator then moves to that location from his current location and then scans the Warehouse Management System (WMS) allocated possible location for it. If the product is not present then he/she scans the aisle (depending on the RFID tags used the scan range may be a few feet or it may be hundreds of feet). If he receives a signal whose strength is limited (there are various thresholds that may classify the signal strength (alternatively, read-rates as strong, weak, intermediate and so on) then he/she moves in a direction that gives him a better signal strength compared to the other directions. These received signal strengths are called observations and, in this model, we have considered five kinds of observations-signal increase, signal decrease, no signal, item found (or maximum signal strength) and no-change. The operator then moves in one of the possible directions which translate into updating the belief states in the POMDP domain (as an action has been taken).

In the WS model each state has five possible categories of observation based on the signal strength received at that location. We have categorized the signals received from querying the tag into the following:

- *'NoSignal'*- Indicates that the forklift operator did not receive any signal
- *'NoChange'*- Indicates an observation in which the operator takes a 'stayput' action and hence receives the same observation as his previous one
- *'SignalDecrease'*- Indicates that the forklift operator has received a signal whose power is less than what he received at the previous location. It is most likely that he has moved away from the tag though this is not certain because of the RFID system disadvantages discussed earlier
- *'SignalIncrease'*- Indicates that the forklift operator has received a signal whose power is more than what he received in the previous location. Here too it is highly probable that he has moved towards the location of the tag but not with complete certainty
- *'ItemFound'*- Indicates that the operator has located the item in the aisle that he is in currently. This observation is only made if the operator has physically scanned and picked up the item from the rack.

Hence the set of observations Ω as used in the representation of POMDP could be represented by

$$\Omega = \{NoSignal,\ NoChange,\ SignalDecrease,\ SignalIncrease,\ ItemFound\}$$

4.5 Reward Model

The maximum reward of 5.0 is for observing (receiving) an *'ItemFound'* signal at locations

$$S = \{(0, 0), (1, 1), (2, 2), (3, 3), (4, 4), (5, 5)\}$$

Since there exists a 'Wall' or a region beyond which the operator is not allowed to search, there is also a penalty associated with trying to reach that location. In the WS model, the Location 6 which is also the boundary for the entire grid is considered a 'Wall' state. The operator incurs a heavy penalty (cost or negative reward). In a real-time warehouse environment, the interpretation would be that searching beyond a fixed distance from the WMS-assigned location (or starting location) is too costly, i.e., the product is unlikely to be found within the allotted time to search for this product. There is a cost if the FLO tries to move to the wall state from each of the locations and it is -10.0.

Note: If the tag is in the same aisle as the operator, he/she gets a 'SignalIncrease' observation with very high probability (the values depending on the observation models discussed later) if he moves towards the tag. A 'SignalDecrease' observation with very high probability value is received when he moves away from the tag in the same aisle. This is because the same aisle has very little or no obstruction for the reader to scan and locate the item. However if he is in an adjacent aisle and moves away from the tag, he gets the ‘SignalDecrease’ or ‘NoSignal’ observations with varying probabilities since cross aisle detection by the reader is limited.

The above note could be well explained with an example (see Figure 6). Consider the operator to be at location 5. Since he is unsure of the location of the tag the initial state could be (5, x) to him. After taking the action LEFT, if he observes *‘SignalIncrease’* with 0.9 probability then he concludes that the tag could be in the same aisle (with more probability that it is in 3, i.e. (5, 3) than it is at 0, i.e. (5, 0)). Using the Bayesian update equation the belief update operation is carried out. However if he observes *‘SignalIncrease’* with 0.5 probability or less then he concludes that the tag could in the adjacent aisle i.e. more probability at (5, 0) than at (5, 3).

To compute the value functions for each horizon length in a POMDP model we have to specify the set of states, actions, rewards (or costs incurred) received due to a particular action and the observations that we get in a particular state (these observations need not necessarily depend on the action).

3	4	5 (FLO)
0	1	2

Figure 6: Aisle configuration for POMDP model

Moving away from the tag incurs costs and the farther the forklift operator moves from the tag the more the costs. Given that he starts in any of the positions with equal probability the operator must navigate to the location of the tag with the help of the POMDP model. Hence based on the above scenario the following descriptions arise:

Since the joint location of the RFID reader and the tag is unknown, the general value iteration algorithm cannot be used to model this situation. What is required is a modification of the value iteration algorithm, which takes into account the uncertainty of the current state. Thus, as mentioned in the brief write up on POMDP we would need to compute the value functions over 'belief states' for the continuous state MDP.

The code for this implementation of POMDP was obtained from www.cassandra.org. It is maintained by Dr. Anthony R. Cassandra and allows us to compile it in a Linux machine. There are command line options to implement any of the following algorithms used to solve POMDPs:

1 Witness
2 Two Pass
3 Linear Support
4 Exhaustive enumeration
5 Incremental Pruning

Incremental Pruning is the algorithm predominantly used to solve the following scenarios since it has been proved to be better than the others in a comparison study by Cassandra et al. [19].

Implementation Details

The format for the input model is specified in the POMDP-SOLVE program. A brief explanation for each input is given below and the detailed input with all states, actions, rewards and observation models are given in the appendix.

1. **Discount Factor γ :** This describes the preference of an agent for current rewards over future rewards. When γ is close to 0, rewards in the distant future are viewed as *insignificant.* We assume a discount factor of 0.95 initially.

2. **States:** The state description for the WS problem is a combination of the location of the forklift operator and the tag and hence in the 5- grid scenario the possible states are 43 in number with (0, 0) to (5, 6) representing 42 states and the 6^{th} state (Wall) representing the 43^{rd} state
3. **Actions**: The five actions as specified above –up, down, left, right, stayput
4. **Observations**: ItemFound, SignalIncrease, SignalDecrease, NoSignal, NoChange
5. **State/Transition model:** It is a matrix which is 43X43 and is input in the format given below:

To specify an entire transition matrix for a *particular action*:

T: <action>	End state 1	End state 2	End state 3…
Start state 0	**%f**	**%f**	**%f…**
Start state 2	**%f**	**%f**	**%f**
Start state 3 …	**%f** …	**%f**	**%f**

Where each row corresponds to one of the start states and each column specifies one of the ending states. The state numbers go from left to right for the ending states and top to bottom for the starting states. The only restriction is there must be NxN values specified where 'N' is the number of states.

6. **Observation Model:**

To specify a row of a particular actions observation probability matrix:

O:<action>:<end-state>

%f %f ... %f

This specifies a probability of observing each possible observation for a particular action and ending state. Since there are 5 actions and 43 end-states, there has to be 43x5 (=215) observation specifications for each model.

7. **Reward/Cost specifications**:

To specify individual rewards:

R:<action>:<start-state>:<end-state>

%f %f ... %f

Please see the appendix for complete input specifications.

CHAPTER V

WAREHOUSE SEARCH MODEL OUTPUT ANALYSIS

To compare and evaluate the difference between a no- RFID system and having one with different signal strength observations (uncertainties) we have considered the effects of three factors as shown in Figure 7.

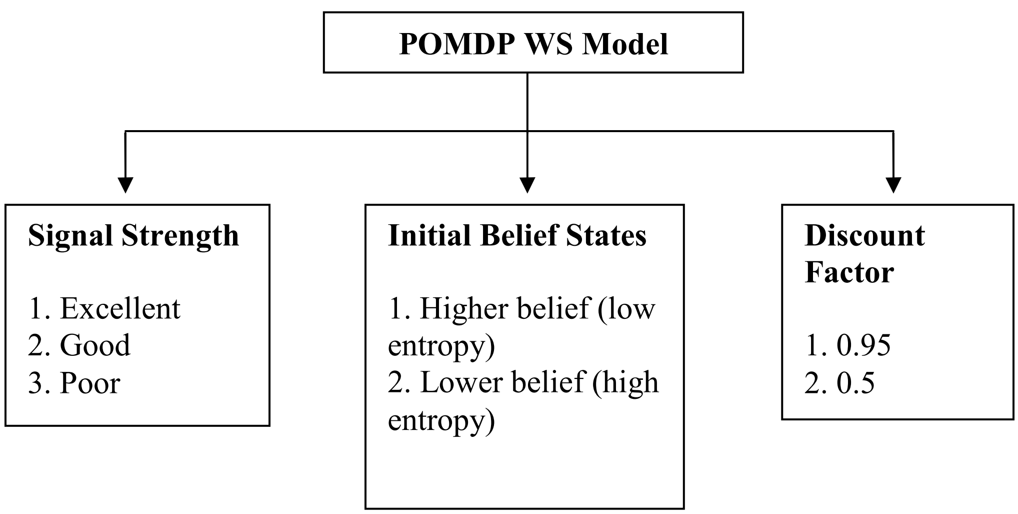

Figure 7: POMDP-WS model categories

5.1 Effect of Signal Strength Observations

Excellent Observations

In this case the forklift operator receives very high fidelity signals throughout the warehouse, i.e., the signal strengths or the power of the received signal is highly indicative of the actual distance between the forklift operator and tag. Such scenarios are likely in the presence of active tags or possibly in a well laid out warehouse that uses passive tags. The observation probabilities are close to 1 for one of the observations and close to zero for all others in this case.

Consider Figure 8 where the tag is assumed to be at Location 5. For this model, the forklift operator receives a 'Signal Increase' observation with complete certainty i.e., probability of one if he moves towards the tag (example, from Location 1 to Location 4). Similarly he receives a 'SignalDecrease' observation with complete certainty if he is moves away from the tag (example, Location 3 to Location 0).

3	4	5 (tag)
0	1	2

Figure 8: Movement of FLO from 1 to 4

Good Observations

This scenario mimics an environment with passive RFID tags. The values of the probability of observation of each element in the observation set are different at different locations, but they are generally not close to 1. If the value of an observation's probability is high compared to some other observation in the set, then it is said to be a 'Good Observation Probability'.

The ‘Good’ signal strength observation case consists of two models. In general, for all the models in the good observation probability scenarios; the operator receives a ‘SignalIncrease’ with much higher probability if he moves towards the tag. However, for this particular scenario- Scenario 1 (GS1); there is also a possibility of a ‘NoSignal’ observation if he tends to be too far away from the tag, and such a no-signal scenario is not considered in Scenario 2 (GS2). GS1 has observation probability values ranging from 0.7 to 0.9 with three possible observations in each case depending on the distance of the tag from the FLO. GS2 has values from 0.8 to 0.95 but with just two possible observations in each location depending on the distance between the tag and the FLO.

To illustrate this further, Figure 9 shows the movement of the operator from Location 1 to Location 4 taking the ‘Up’ action. Since the tag is at 3, the FLO receives a ‘SignalIncrease’ observation with probability 0.8 and a ‘SignalDecrease’ observation with a probability of 0.2. In this case he does not receive a ‘NoSignal’ observation because he is in the same aisle as the tag and has more chances of receiving a correct indication.

3 (tag)	4 (FLO)	5
0	1	2

6

Figure 9: Up action by operator

To distinguish the effects of the observation probabilities, in Scenario GS2, the FLO receives a ‘SignalDecrease’ with an extremely high probability value if he moves away from the tag. This is because, in general, when the distance between the source and destination increases, the signal strength decreases. We have negated the possibility of a 'SignalIncrease' to represent a better signal strength observation scenario.

In Figure 10 the FLO moves 'Up' from Location 2 to Location 5 and the tag is at Location 3. Table 1 following Figure 10 shows the observation probability values in this case compared to the values for a similar scenario in the previous case (GS1).

. IF corresponds to 'ItemFound', SI- 'SignalIncrease', SD-'SignalDecrease', NS- 'NoSignal' and NC- 'NoChange'.

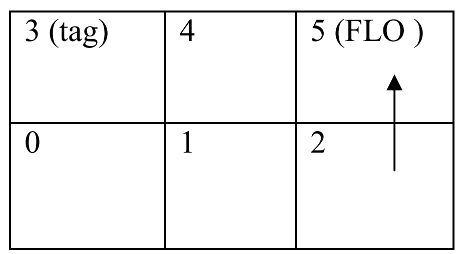

Figure 10: Good observation scenario (GS2)

It can be noticed from Table 1 that GS2 has less amount of uncertainty when compared to GS1. Similar to the example shown above, the other values of observation probabilities are constructed to reflect the case. Specifically, in the third column the 0.1 probability of the no signal observation increases the uncertainty.

Table 1 Comparison of observation probabilities

	IF	SI	SD	NS	NC
scenario 1	0.0	0.8	0.1	0.1	0.0
scenario 2	0.0	0.8	0.2	0.0	0.0

Please refer to the appendix for complete input specifications for all the models discussed above as well as the poor observation model discussed below.

Poor Observations

A 'Poor (Bad) Observation Probability' is one in which the operator receives equi-probable signals. Assume that the operator moves from location 3 to location 4 and receives a 'SignalIncrease' with probability 0.5 and a 'SignalDecrease' with probability 0.5. This does not help us to update the belief state appropriately. This scenario is akin to the one that does not use reliable information.

The effects of these observations are studied in terms of the policies and rewards in the future sections. The output files of a POMDP program are a set of alpha vectors and a policy graph with extensions .alpha and .PG. The .alpha file is organized in the following way:

Table 2 Format of output .alpha file

<**vector 0-action**>
<vector 0-coefficient-0> <vector 0-coefficient-1> ... <vector 0-coefficient-n-1>
<**vector 1-action**>
<vector 1-coefficient-0> <vector 1-coefficient-1> ... <vector 1-coefficient-n-1>
….

This is a representation of the hyperplanes, whose upper surface (the maximum over the belief space) defines the value function for the problem. The first line is the action that is associated with each facet of the surface and then the coefficients for the hyperplanes are given. Each component of the vectors has this rough interpretation:

- Each vector, aside from its coefficient values, has an associated action a

- Component $'i'$ of a vector says that if we were certain we were in state $'i'$, and took action a, and followed the optimal policy thereafter, then we would receive this much reward in the infinite, discounted horizon
- Then we multiply this value by our belief of being in state $'i'$, and repeat for all other components to get an expected reward for the given belief state if we took action a and then followed the optimal policy after that

5.2 Effect of Initial Belief States

There are two types of initial or starting belief states that the operator could have depending on the WMS information available or scanning from the current location. As suggested in Figure 7, a good initial belief is one in which the operator is more certain as to how far the tag could be from his current location. If the forklift operator is sure that the tag is within the read range of the reader and is close by, his/her belief state has a focused probability distribution.

5.2.1 Higher Belief Probabilities (Good Initial Belief States or Lower Entropy States)

If the forklift operator is sure that the tag is within the read range of the reader and is close by, his/her belief state has higher probability values. Such a scenario would arise in the presence of a WMS in a well managed warehouse. Here the item is likely to have moved to a location in the close proximity of the location suggested by the WMS. An example would be to believe that the tag is at 3 more than the tag is at 0 (see Figure 11).

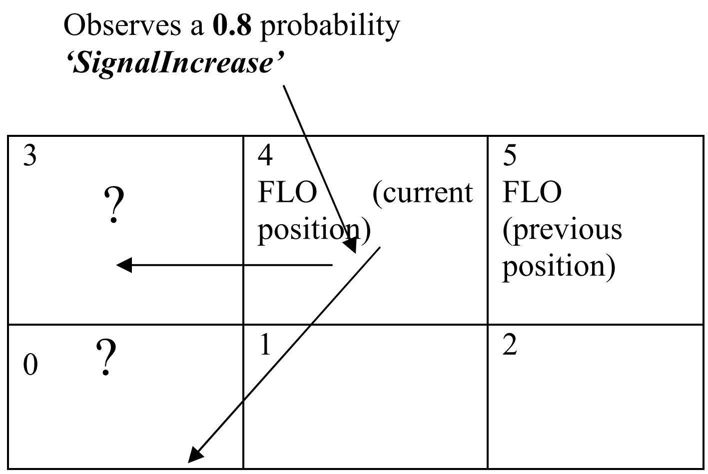

Figure 11 Example of a higher belief state (lower entropy) condition

Mathematically, the belief space would be represented as follows (assuming that the operator is at location 4 in the grid)

$$b\{s = (4,3)\} = 0.8, b\{s = (4,0)\} = 0.2$$
$$b(s) = 0 \forall s \notin \{(4,3), (4,0)\}$$

Assume that the forklift operator is assigned to some location. If the item was moved from that location and misplaced within the read range of the forklift operator, he is sure to get some signal at his current location. But if he does not get a signal it could mean either or both of the following could have happened:

- The item was moved to a place which is very far from the read range and hence is undetectable by the forklift reader
- The item is still within the read range of the reader but the effects of path loss or metals or liquids or a combination of these factors has led to no signal being detected at his current location

This means that he has to take *some* action first to see if he gets a better signal or some signal to indicate that the tag is within the read range. So the *first* action could be random.

5.2.2 Lower Belief Probabilities (Bad Initial Belief States or Higher Entropy States)

When the forklift operator receives a 'SignalDecrease' observation or a 'NoSignal' observation, he is not sure of the location of the tag. If the operator reaches 4 from some other location (see Figure 12) in the warehouse and receives a *'SignalDecrease'* or *'NoSignal'* observations (the tag is only partially observable) he is very unsure as to where the tag is. Though he is more certain that the tag cannot be in the same aisle he is unsure as to which rack the tag is in the adjacent aisle. Mathematically, his current belief state is now divided equally between locations 0, 1 and 2.

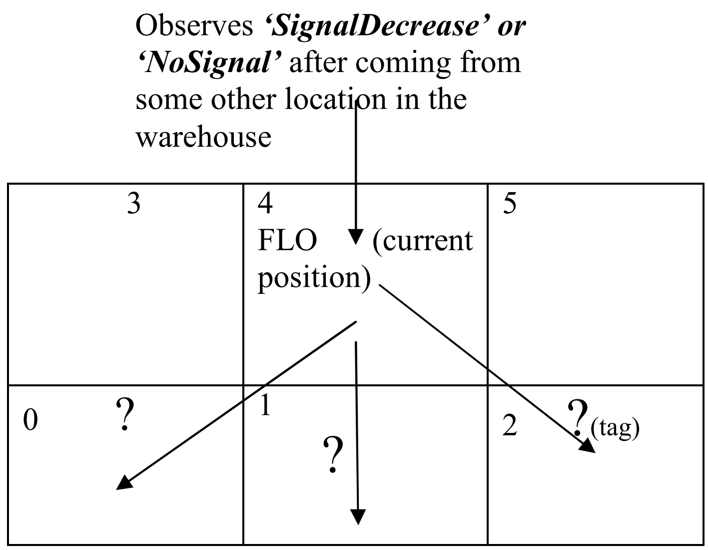

Figure 12 Example of a lower belief state (higher entropy) condition

In this case let us assume that the forklift operator is at location 4 and the initial belief states are equally biased between location 0, 1 and 2. The probability that the tag is at location 0 is the same as the belief probability the tag is at 2 and at 1 represented mathematically by

$$b\{s=(4,0)\}=(1/3), b\{s=(4,1)\}=(1/3), b\{s=(4,2)\}=(1/3)$$
$$b(s)=0\,\forall s\notin\{(4,0),(4,1),(4,2)\}$$

When the agent (the operator in the WS problem) is very uncertain about the real underlying state of the world, it is said to be in a belief state of high entropy. In such belief states, the agent cannot select actions very appropriately [11].

5.3 Quantitative analysis of results- comparison of expected future reward

All the models below (Section 5.3.1 to Section 5.5.3) assume that the *forklift operator starts from location 0*. The model is built in such a way that the starting state has no influence on the rewards or costs accrued to ensure generality. Also it is to be noted that the algorithm produced vectors only when the ε- epsilon (the precision with which the pruning operations occur) value was closer to 30 or above (for the tested scenarios). Hence, for all the five different scenarios (excellent, GS1, GS2, poor, no-RFID) the epsilon value has been fixed to be 30.

In addition, to compare the expected discounted reward obtained in each of the scenarios we have to have a common base model of a belief state. This is very important because different initial belief states lead to different policies (action-observation sequences) and hence different rewards are accrued for each of the scenarios. We aim to explain the differences in the reward accrued *based on the reliability of the observations received*. Excellent observations are of the highest reliability as to what state the system is in currently and, good observations are less reliable than their excellent counterparts. The effects of poor observations that are in the 0.5 range are also calculated for the same initial belief state as with the other cases.

Let the current belief state of the operator be represented as follows:

$$b\{s = (0,2)\} = 0.8, b\{s = (0,1)\} = 0.2$$
$$b(s) = 0 \forall s \notin \{(0,1), (0,2)\}$$

The following sections compare and contrast the movements of the forklift operator in terms of the actions suggested and numerical rewards obtained for different decision epochs for the different cases of the POMDP-WS model as described in Figure 7.

5.3.1 Excellent Observation Probabilities

In this model the observation probabilities occur with probability one. This means that if the FLO is moving towards the tag, he always gets a signal increase (and no other observation). A perfect RFID system is a good example of this model. If he is moving away from the tag (whichever aisle he is in), he gets a decreased signal strength. As stated above, it is a deterministic case in which the observations obey the Friis transmission equation. Table 3 captures the expected future reward obtained for each decision epoch. We calculate this by using the alpha vectors of the value functions generated using the incremental pruning algorithm. Multiplying the belief state with the alpha vectors helps us in identifying the starting node in the policy graph file. Since the result of the multiplication is a single number for each vector, we find the maximum of it and that corresponds to the starting node in the policy graph. The corresponding action suggested is the first optimal action to take that also gives the maximum reward as it corresponds to the best action for that belief state at that decision epoch.

Table 3: Excellent Observation Probabilities

Epoch	**Suggested First Action**	**Expected Future Reward**	**Number of nodes in PG* file (Number of vectors)**
1	stayput	0.00	1
2	right	0.95	5
3	right	5.46	10
4	right, right[1]	9.74	11
5	right, right	13.81	13
10	right, right, right	31.32	15
15	right, right, right	44.87	15
20	right, right, right	55.35	15
508 (∞ horizon) converged	right, right, right	91.19	15

[1] We could start at any of the two nodes suggested as both correspond to the same starting state. What it means is that the expected future reward is numerically equal for these two nodes and the FLO is free to choose any of them to make a move. This is also true for cases in which there are three starting nodes.

* If the numerical values of the maximum expected reward are equal for two different actions, the operator can choose any one and proceed to the next step. Choosing a particular action takes the operator through a specific route in the policy graph.

PG is the policy graph file that is generated with the alpha vector file. This contains the policy graph to be used in the case of the converged infinite horizon solution. In the case of a finite horizon problem it suggests the first action to take to the forklift operator.

5.3.2 Good Observation Probabilities

Table 4 summarizes the results of good observation probabilities- GS2 in a warehouse in the 0.8 range. This model did not converge to a solution as a policy graph for a discount factor of 0.95. Therefore this model maybe used only for providing short epoch (<20) policies.

Table 4: Good Observation Probabilities-GS2

Epoch	**Suggested First Action**	**Expected Future Reward**	**Number of nodes in PG file (Number of vectors)**
1	stayput	0	1
2	right	0.90	5
3	right	5.35	11
4	right, right	9.66	14
5	right, right	13.70	15

10	right	31.03	16
15	right, right, right	44.32	15
20	right	55.03	16
508 (did not converge)	right	90.14	16

5.3.3 Bad/Poor Observation Probabilities

In this model the observation probabilities are equally divided, i.e., if FLO moves towards the tag he gets a signal increase with a probability of 0.5 and a signal decrease with a probability of 0.5. Thus, it is extremely confusing as to whether he is moving towards the tag or away from the tag at any location. The initial belief state is same as the one for the good and excellent observation case. Table 5 summarizes the results of the poor observation case and it can be clearly seen that the expected reward is much lower than that for the other two cases described earlier.

Table 5: Bad Observation Probabilities

Epoch	**Suggested First Action**	**Expected Future Reward**	**Number of nodes in PG file (Number of vectors)**
1	stayput	0.00	1
2	right	0.5	5
3	right	1.25	11

4	right, right	2.37	14
5	right, right	2.68	15
10	right, right, right	2.98	15
15	right, right, right	2.98	15
20	right, right, right	2.98	15
35 (converged)	right, right, right	2.98	15

5.3.4 Comparison of expected discounted rewards with good initial belief states

Figure 13 captures the expected future reward for a good initial belief state model with different values of signal strength observation probabilities.

It can be seen that if the forklift operator has a very accurate and focused initial belief state then the signal strength observations are less significant. This is also true for the no-RFID signal case because he may be guided by accurate information from the warehouse management system. However, the poor observation scenario has a very small reward of 2.98. This is because in this model the forklift operator may reach locations that are undesirable even with better belief states. There is also a very small difference in the reward values for GS1 and GS2.

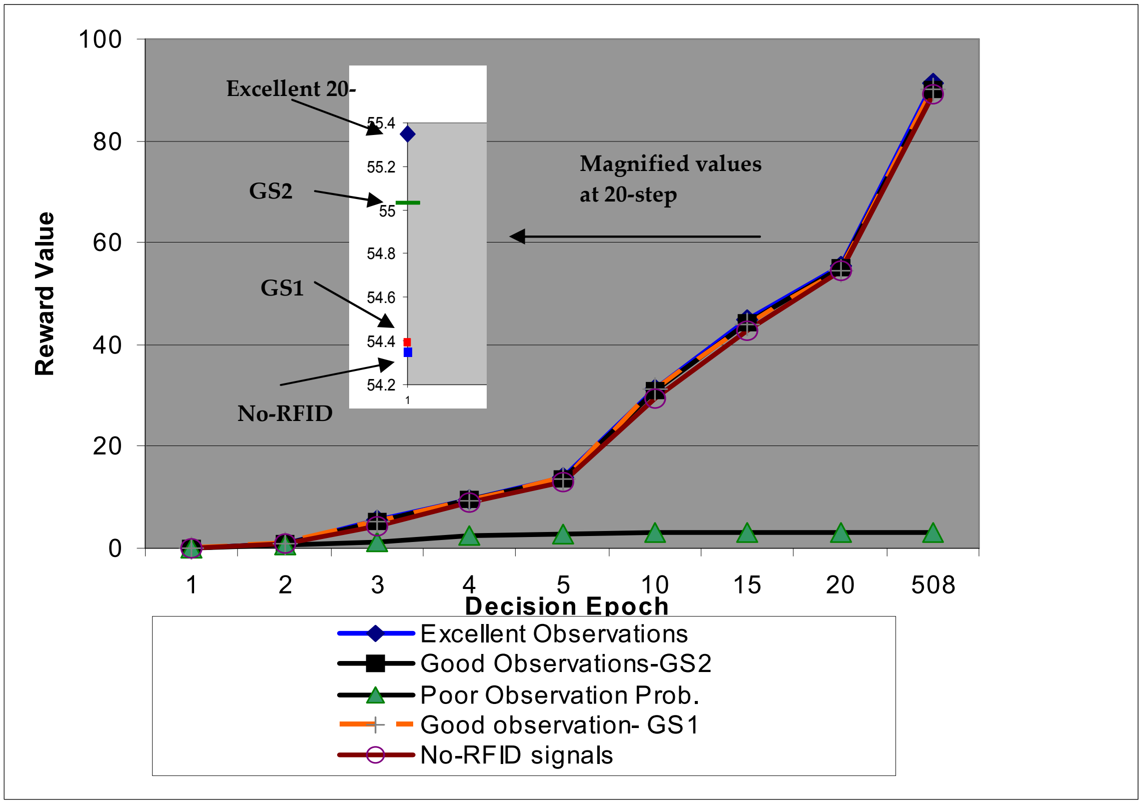

Figure 13: Comparison of rewards for different observation scenarios

5.4 Comparison of expected rewards with bad initial belief states

Assuming that the forklift operator reaches a location where he does not receive any signal, the belief state at such a location is said to be bad. Consider the case when the operator reaches Location 0 and does not receive any signal. Then his current belief state could be represented by the following expression

$$b\{s=(0,1)\}=1/5, b\{s=(0,2)\}=1/5, b\{s=(0,3)\}=1/5, b\{s=(0,4)\}=1/5, b\{s=(0,5)=1/5\}$$
$$b(s)=0 \forall s \notin \{(0,1),(0,2),(0,3),(0,4),(0,5)\}$$

Since he does not know the exact location of the tagged item, he believes that the tag could be anywhere in the grid except at his current starting position 0. Thus, for a discount factor of 0.95, the following sections explore the impact of different observation models starting with the no RFID case in Section 5.4.1

5.4.1 No RFID signals

This case represents a warehouse with no RFID system information. The only observation is ‘NoSignal’ that is received at all the locations except at ones where the tag could actually be. The following locations or states in the WS model have the ‘ItemFound’ observation:

$$S = \{(0, 0), (1, 1), (2, 2), (3, 3), (4, 4), (5, 5)\}$$

This is because the forklift operator can see the item only when he is in the same rack as that of the tag. At all other locations he gets a ‘NoSignal’ observation. On running this model, we found that it converged producing 17 vectors. The expected discounted reward for this scenario was calculated to be 66.68.

5.4.2 Excellent, Good and Poor observation probabilities

The expected reward comparing the no-RFID case with excellent, good and poor observation probabilities is shown as a graph in Figure 14 below.

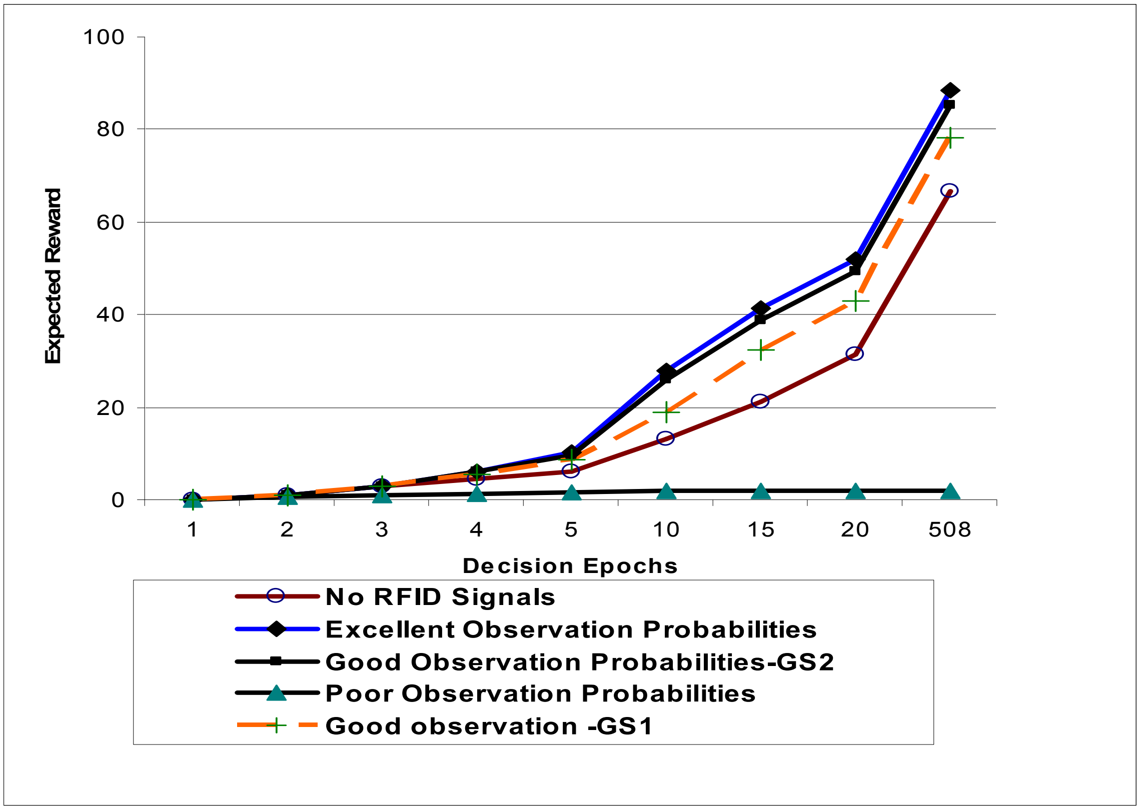

Figure 14: Comparison of expected rewards for bad initial belief states

Figure 14 captures the expected future reward for a bad initial belief state model with different values of signal strength observation probabilities. From the graph the following can be inferred:

- The excellent observation probability scenario has a very high expected reward for even the bad initial belief state (88.62)
- The good observation probability GS2 scenario has an average reward that is slightly less (85.10) compared to the excellent case
- The good observation probability scenario GS1 has an average reward that is less (78.23) compared to GS2
- The no-RFID case has a reward of 66.68 which is 56% lower than that for the good observation cases

- The poor observation probability model has the least expected discounted reward of 1.87

The main result here is that the expected reward for GS2 is about 56% higher than that of the no-RFID scenario. This means that the forklift operator can locate the item expending significantly less time and efforts even with imperfect RFID signals compared to a scenario having no RFID system installed. An additional take away from this investigation is that the no-RFID signal case fares much better in terms of expected reward compared to the poor observation scenario (which is less probably in a well designed RFID system). This is because in the poor observation scenario, the forklift operator is significantly misguided by the false observations received. If there was no RFID system in place he is better off searching the locations one by one rather than having to come back to the same location as suggested by the poor observation model.

Table 6: Comparison of Expected reward values for bad initial belief states (with 0,95 discount factor)

Excellent obs.	**Good Obs. GS1**	**Good Obs. GS2**	**No-RFID**	**Poor signal obs.**
0	0	0	0	0
0.95	0.95	0.95	0.95	0.5
2.75	2.84	2.75	2.754	1.0
6.178	5.32	6	4.46	1.4
10.25	8.58	9.668	6.098	1.65

27.75	19.01	25.814	13.09	1.87
41.3	32.44	38.81	21.04	1.87
51.78	42.86	49.316	31.52	1.87
88.62	78.23	85.106	66.68	1.87

5.5 Effect of discount factor

The scenarios discussed above were executed with a discount factor of 0.95. If the discount factor is very high (close to 1) then the future rewards have more effect on current decision-making. However, we also evaluated scenarios by changing the discount factor to 0.5. It is very interesting to note that all the models with discount factor 0.5 converged in a very short time period (a few seconds) and hence, we were able to get a complete policy graph for each of them for the infinite horizon case.

Decreasing the discount factor from 0.95 to 0.5 decreased the long term rewards considerably but even if the reward values are numerically equal, the actions suggested by the POMDP program differ and hence the number of steps to reach the tag differs based on the starting belief state of the operator. Figure 15 shows the expected reward values for a 20-step horizon for all the models. The maximum reward for the 0.95 discount factor obtained by the operator in the excellent observation case is 51.78, in GS1 42.86, GS2 with less uncertainty 49.31, poor observation case 1.87 and no-RFID signal case 31.52. Comparatively, for the models with discount factor 0.5, the rewards for each of the 5 cases were 2.116, 1.93, 2.054, 1.87, and 1.49, respectively.

It is interesting to note here that the No-RFID case has a reward of 1.49 which is almost 25.5% lesser than that of the poor signal strength case (1.87). As the rewards of the future steps are increasingly discounted, the expected reward of the poor signal strength case is more than the no-RFID signal case.

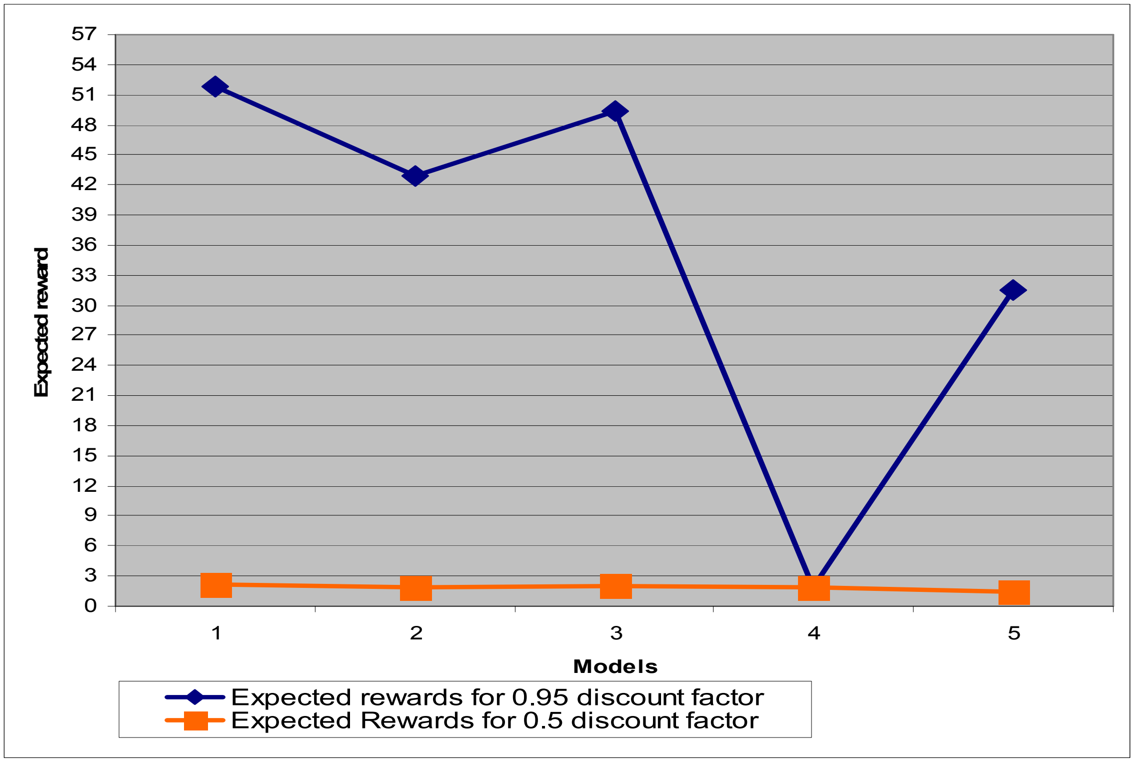

Figure 15: Comparison of 20-step expected rewards for discount factors 0.95 and 0.5

5.5.1 Effect of bad belief states and excellent observations

If the forklift operator reaches Location 0 and his initial belief state is bad i.e., he believes that the tag could be present either at Locations 1, 3 or 4, then the current belief state is represented by

$$b\{s=(0,1)\}=1/3, b\{s=(0,3)\}=1/3, b\{s=(0,4)\}=1/3$$
$$b(s)=0 \forall s \notin \{(0,1),(0,3),(0,4)\}$$

The POMDP algorithm for the excellent observation model with a discount factor of 0.5 suggested two actions- 'right' or 'up' that had numerically equal expected rewards of 2.99 for this belief state. Two scenarios one for each action has been described below.

Figure 16 shows the movement of the forklift operator starting with action ‘right’. After taking the action ‘right’, the operator enters Location 1 from Location 0. If the tag is at Location 3, then he observes a ‘SignalDecrease’ observation at Location 1. This is because this is a model that assumes a perfect RFID system.

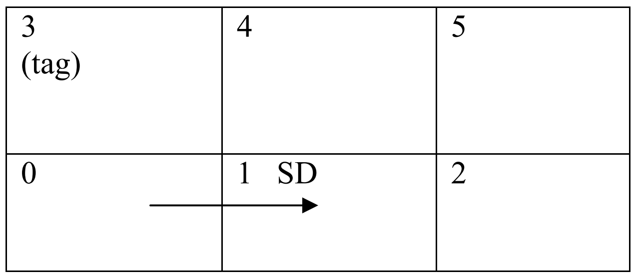

Figure 16: FLO movement with excellent observation probabilities

The action suggested at Location 1 after the ‘SignalDecrease’ observation is ‘left’. This takes the operator back to the starting Location 0. Thus, he skips scanning Location 2. At Location 0, he receives a ‘SignalIncrease’ observation. This is because he has moved towards the tag at Location 3. This is clearly indicated in Figure 17.

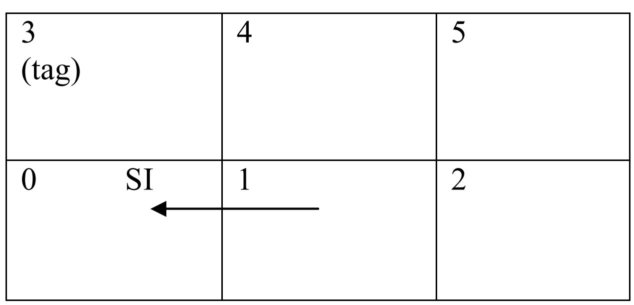

Figure 17: Movement of FLO based on excellent observations

The next action suggested at Location 0 is ‘up’. This takes the operator to Location 3 where the ‘ItemFound’ observation is received. Here, the final action suggested is ‘stayput’ which makes the operator to stay at Location 3.

In the event that the tag is at Location 4, then the observation received at Location 1 is a ‘SignalIncrease’. In this case, the algorithm suggests taking an ‘up’ action at Location 1. This takes the operator to Location 4, where he receives the ‘ItemFound’ observation.

If the operator decides to choose the action ‘up’ (starting node 7 in the policy graph file) action instead of ‘right’ at Location 0, then it takes him to Location 3. If the tag was at 4, he gets a ‘SignalIncrease’ observation at Location 3. After updating the belief state with this observation, he takes the action ‘right’ suggested in the policy graph. The ‘ItemFound’ observation is received at Location 4 since the tag is located there.

5.5.2 Effect of bad belief states and good observations

In the case of good observations the forklift operator receives ‘SignalIncrease’ and ‘SignalDecrease’ or ‘NoSignal’ in different probabilities varying from 70 to 90%. In these cases the forklift operator takes more steps to reach the tag as discussed in the case below.

Let the operator have an initial belief state as represented by the following equation when he is at Location 0

$$b\{s=(0,1)\}=1/5, b\{s=(0,2)\}=1/5, b\{s=(0,3)\}=1/5, b\{s=(0,4)\}=1/5, b\{s=(0,5)=1/5\}$$
$$b(s)=0 \forall s \notin \{(0,1),(0,2),(0,3),(0,4),(0,5)\}$$

In this case the POMDP algorithm suggests taking the action ‘right’. This is shown in Figure 18. Let the tag be at Location 4. In spite of the observation here at Location 1 (‘SignalIncrease’ or a ‘SignalDecrease’) the algorithm suggests taking action ‘right’. Since he is moving *away* from the tag at Location 4, he must receive a ‘SignalDecrease’. The policy graph actually suggests the same action for both the observations. This is denoted by the * in Figure 18. It suggests an action ‘up’ that takes him to Location 5 in the grid. At Location 5 a ‘left’ action is suggested and at Location 4 a ‘down’ action is suggested irrespective of the observation received.

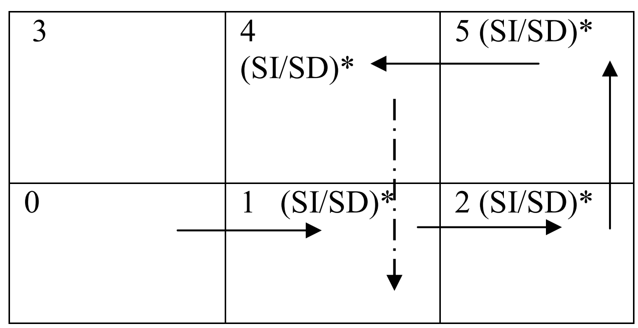

Figure 18: Effect of bad belief states and good observations

5.5.3 Effect of bad belief states and poor observations

Figure 19 shows the movement of the forklift operator for the poor observation model. In this model the observations are equi-probable i.e., the probability of a ‘SignalIncrease’ is the same as that for a ‘SignalDecrease’. This leads to poor observations and hence poor updated belief states.

Let the operator have an initial belief state as represented by the following expression when he is at Location 0

$$b\{s=(0,1)\}=1/5, b\{s=(0,2)\}=1/5, b\{s=(0,3)\}=1/5, b\{s=(0,4)\}=1/5, b\{s=(0,5)=1/5\}$$
$$b(s)=0 \forall s \notin \{(0,1),(0,2),(0,3),(0,4),(0,5)\}$$

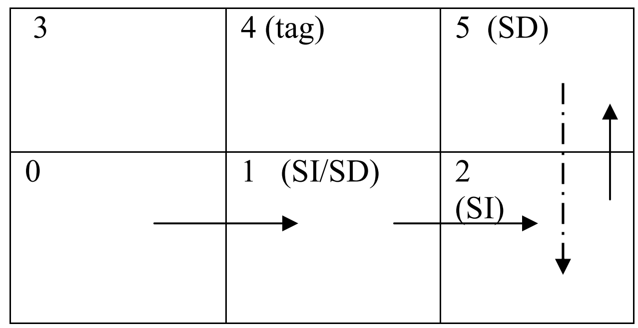

Figure 19: Effect of bad belief states and poor observations

Similar to the good observation scenario, the policy graph in the poor observation scenario suggests ‘right’ as the first action to be taken. This takes the operator to Location 1. Irrespective of a ‘SignalIncrease’ or a ‘SignalDecrease’ observation at 1, the action suggested is ‘right’. This action takes the operator to Location 2. If the tag is assumed to be at Location 4, then the observation at Location 2 should be a ‘SignalDecrease’. However, this being a poor observation model, the observation maybe a ‘SignalIncrease’. In such a case, the algorithm suggests taking action ‘up’. This takes the operator to Location 5 (he would move back to Location 1 if this was an excellent observation model).

The poor observation model is such that there is a 50% chance of observing a ‘SignalDecrease’ at Location 5 even if the tag is at Location 4 (very close-by). If the observation is a ‘SignalDecrease’, the operator is forced to take the action ‘down’ which takes him to Location 2 thus, revisiting the location again. This prevents him from reaching the tag’s location and sometimes might not take him to Location 4 at all. This scenario is highly undesirable.

5.6 Discussion of Results

The excellent observation model is undoubtedly the best in that it consistently provides, under good initial beliefs, the shortest path to the forklift operator to reach the tag both in terms of expected rewards and policies. The good observation scenario GS1 with more uncertainty in the signal strength observations has rewards lesser than GS2 where the uncertainty is reduced. Even with that uncertainty built into the model the operator is able to choose actions depending on observations and reach the tag with more number of steps compared to the excellent observation case.

However, in a poor observation model, the FLO gets to the tag faster if he receives the right observations (if he is lucky), but may never reach the tag if the observations are poor. This is because he is forced to change the course of his direction because he receives a wrong observation which makes him revisit previous locations.

Also, in terms of the expected reward values for 20- time step decision horizon, it can be seen that with good initial belief state the excellent observation case has a 18 times numerically higher reward than the poor observation case. With the infinite horizon case, the difference between the rewards for the excellent observation case and the poor case is almost 30.

As the observation probabilities are decreased i.e., there is more randomness in the nature of signals received (observations), the number of steps to reach the tag increases considerably. The expected reward from a 20-step POMDP with reasonable observation probabilities (varying between 70 and 90%) was 18 times higher than for the poor observation model. This result is tantamount to a ~18 time reduction in search times and efforts for locating a misplaced item using RFID in the tested scenarios.

The good observation model one has more uncertainty with all three observations of 'SignalIncrease', 'SignalDecrease' and 'NoSignal' being non-zero. Comparing the expected reward for a 0.95 discount factor and bad initial belief states we find that model two has a 9% better reward numerically. Decreasing the discount factor to 0.5 decreases the difference between them to 6.5%.

The no-RFID case model has a reward of 31.52 for the 20-step horizon, which is almost 56% lesser compared to the good observation model two. It is also 36% lesser compared to the good observation scenario (GS1) with more uncertainty. This result is tantamount to a ~56 % reduction in search times and efforts for locating a misplaced item using RFID in the tested scenarios. This is significant as it points to the possible benefit from the use of RFID despite its imperfect information. Also the no-RFID case has a much better expected reward (66.68) in the infinite horizon with discount factor 0.95 compared to the poor RFID signal case. However, with a discount factor of 0.5, the no-RFID case has a reward which is 25.5% times lesser than the poor signal strength case.

CHAPTER VI

CONTRIBUTIONS AND FUTURE WORK

6.1 Contributions

In this research we have investigated the value of using RFID signals to make the searching process for misplaced items more efficient. Missing items are known to account for a very large percentage of lost sales as observed by Ton and Raman [1]. Companies are known to spend significant amount of times searching for misplaced items. Given these disadvantages, a search process using RFID signals as beacons to identify the misplaced items efficiently has been formulated using POMDP. We have found that the expected value (reward) of the search process with excellent signal observations is undoubtedly the best in terms of numerical rewards as it is 18 times greater (in the case of 20-step reward) than the case in which the observations (signal strengths) are bad for the specific reward model we have used. It has also been found that even with imperfect RFID information one can see a significant value addition compared to a search process without RFID.

In the case of excellent observations, the effect of initial belief states is negligible and the FLO is able to reach the tag based on observations alone. However in the case of the good observation probabilities and bad belief states, the FLO tends to scan some locations while he skips some others which maybe infeasible.

Decreasing the discount factor and considering only near future rewards while heavily discounting off the late rewards makes the good observation case converge. We are able to use the policy graph to help the forklift operator navigate based on initial belief states and observations.

Also, in terms of the expected reward values for 20- time step decision horizon, it can be seen that with good initial belief state the excellent observation case has a 18 times numerically higher reward than the poor observation case. With the infinite horizon case, the difference between the rewards for the excellent observation case and the poor case is almost 30.

As the observation probabilities are decreased i.e., there is more randomness in the nature of signals received (observations), the number of steps to reach the tag increases considerably. The expected reward from a 20-step POMDP with reasonable observation probabilities (varying between 70 and 90%) was 18 times higher than for the poor observation model. This result is tantamount to a ~18 time reduction in search times and efforts for locating a misplaced item using RFID in the tested scenarios.

The good observation scenario GS1 has more uncertainty with all three observations of 'SignalIncrease', 'SignalDecrease' and 'NoSignal' being non-zero. Comparing the expected reward for a 0.95 discount factor and bad initial belief states we find that model GS2 has a 9% better reward numerically. Decreasing the discount factor to 0.5 decreases the difference between them to 6.5%.

The no-RFID case model has a reward of 31.52 for the 20-step horizon, which is almost 56% lesser compared to the good observation model two. It is also 36% lesser compared to the good observation scenario (GS1) with more uncertainty. This result is tantamount to a ~56 % reduction in search times and efforts for locating a misplaced item using RFID in the tested scenarios. This is significant as it points to the possible benefit from the use of RFID despite its imperfect information. Also the no-RFID case has a much better expected reward (66.68) in the infinite horizon with discount factor 0.95 compared to the poor RFID signal case. However, with a discount factor of 0.5, the no-RFID case has a reward which is 25.5% times lesser than the poor signal strength case.

Modeling the uncertainty in the read rates is challenging and hugely depends on the items that warehouses and distribution centers handle. If a warehouse handles a large volume of liquid and metal products and invests in passive tags, this may not be a feasible solution for the searching process. However, due to the ever-developing nature of this technology these difficulties might be overcome in the near future, or, other tag-types may be used which may have excellent read rates in all kinds of environment. Active RFID tags are currently expensive and may not be affordable for a small/mid sized company unless there is a possibility to re-use the tags (closed loop supply chains). This might lead to additional complications like more read/write operations and, affixing and removing the tags at each dock door. The company has to conduct a cost-benefit analysis to see if active tags are affordable and can be reused. If the warehouse deals with costly yet item-level products this may be a very feasible solution given the cost of the product and other factors to locate it quickly.

On the technical side, a lot of tests can be conducted using inputs from actual observations of read-rates from warehouse data from specified locations. Modeling the operation and actions to be taken at each location in the warehouse can be a very dynamic process and can be demonstrated during a pilot-study period. If active tags are used, the efficiency of the search process improves considerably as a very wide area is covered by the reader and the active tag.

6.2 Future Work

In the present study, we have modeled a very small portion of a warehouse to identify if this search process is feasible and what observation probabilities are better from an algorithmic standpoint. The actual warehouse locations could be modeled as regions (instead of the grid) or each region could be split into grid-like structures and this algorithm could be run using inputs from that region. As in any POMDP model, the finite state space and the size of the input state space is a binding factor in modeling a situation with uncertainty. A possible solution to this may be to split the state space of the warehouse. For example, grid Location 4 could be an entire region in itself with 5 aisles and 10 racks. So zooming in on region Location 4 would enable us to look at the problem in a similar perspective as we have considered and modeled in this paper. After reaching any border location in location 4, the next step would be to move to region 3 or 5 or 1 based on where the operator is located currently.

Searching for multiple items as in a regular order-picking operation but using RFID is an interesting challenge that we are currently pursuing. However, the true value of RFID in the search process would only be realized if the entire order-picking operation is considered. Another important aspect of our future research is the complex layouts of warehouses reflecting actual distances traveled. In incorporating this into our model we aim to find the expected travel time taken by the forklift operator to search and track the misplaced items. By calculating the travel times (or the miles traveled by the forklift operator) we can compare and provide a solid model that demonstrates the true value of implementing RFID in a storage environment.

REFERENCES

[1] Z. Ton and A. Raman, "The Effect of Product Variety and Inventory Levels on Misplaced Products at Retail Stores: A Longitudinal Case Study," Harvard Business School, Boston, MA 2004.

[2] E. Sahin, Y. Dallery, and S. B. Gershwin, "Performance evaluation of a traceability system: An application to Radio Frequency IDentification technology," presented at IEEE SMC, 2002.

[3] S.Kirsner, "Building a radar for everyday products," in *Newsweek*, March 2002.

[4] R.B.Brooks and L.W.Wilson, *Inventory records accuracy: Unleashing the power of cycle counting*. New York: John Wiley & Sons, 1995.

[5] Y.Kang and S.B.Gershwin, "Information inaccuracy in inventory systems:Stock loss and Stock out," vol. Ph.D. Boston: Massachussetts Institute of Technology, 2004.

[6] A.Raman, N.DeHoratius, and Z.Ton, "Execution: The missing link in retail operations," in *California Management Review*, vol. 43(3), 2001, pp. 136-152.

[7] "Shrinkage: Introducing a collaborative approach to reducing stock loss in the supply chain," ECR Europe, Brussels, Belguim 2001.

[8] E. Frazelle, *World-class warehousing and material handling*: McGraw-Hill, 2002.

[9] V.Zeimpekis and G.M.Giaglis, "Agile inventory management incorporating wireless technology: warehouse scenario and network topology," Athens University of Economics and Business, Athens 2002.

[10] M. Puterman, *Markov Decision Processes : Discrete Stochastic Dynamic Programming* New York: John Wiley & Sons, 1994.

[11] L.Kaelbling, M. Littman, and A.Cassandra, "Planning and Acting in Partially Observable Stochastic Domains," *Artificial Intelligence*, vol. 101, pp. 99-134, 1998.

[12] J. Brusey, C. Floerkemeier, and M. Harrison, "Reasoning about uncertainty in location identification with Auto-ID," University of Cambridge, Cambridge 2003.

[13] Y.M.Lee, F.Cheung, and Y.T.Leung, "Exploring the Impact of RFID on Supply Chain Dynamics," presented at Winter Simulation Conference, 2004.

[14] H.Lee and O.Ozer, "Unlocking the Value of RFID," Stanford University 2005.

[15] H.K.H.Chow, K.L.Choy, W.B.Lee, and K.C.Lau, "Design of a RFID-based resource management system for warehouse operations," presented at IEEE International Conference on Industrial Informatics, 2005.

[16] A.R.Cassandra, "Exact and Approximate Algorithms for Partially Observable Markov Decision Process," vol. PhD Dissertation Providence: Brown University, 1998, pp. 447.

[17] www.RFIDjournal.com, "RFID Glossary ".

[18] J. Govardhan, "Performance Modeling and Design of Backscatter RFID Systems: A Statistical Approach," in *Industrial Engineering and Management*, vol. Master of Science. Stillwater: Oklahoma State University, 2006, pp. 102.

[19] A. Cassandra, M. Littman, and N.L.Zhang, "Incremental Pruning: A Simple, Fast, Exact Method for Partially Observable Markov Decision Processes," presented at Thirteenth Conference on Uncertainty in Artificial Intelligence Providence, RI, 1997.

APPENDIX

Modeling the effect of inventory inaccuracy

This appendix presents our related investigation on the effect of bad read rates on the accuracy of inventory counts using an Arena® model. An Arena® model is developed to investigate the minimum required read rates to detect the misplaced items in Section A.1. The results of this implementation are also described here.

A.1 RFID read rates for effective item search- A simulation study

Two types of products represented by R and B enter a warehouse based on a constant input distribution (CONST (30) and CONST (35) expression in ARENA) with 30 and 35 pallets of each type arriving separately. They are assigned to racks by the WMS based on some information from the manufacturer. R entities are assigned certain racks and B entities are assigned some other racks in the aisles. Each rack has a capacity of 10 pallets. Because of human error and other uncertainties in the warehouse there is a percentage of each type of entity being placed in the rack meant for the other type. For e.g. R could go into a rack meant for B and vice versa. They stay there for a fixed number of hours before being loaded into the forklifts to be shipped to the retailer according to the demand generated.

A.1.1 Arena® Simulation Model: without RFID and with misplaced items

A simple warehouse environment with 2 aisles and 8 racks in each of them was created as sub models in Arena®. The Arena® model shown in Figure A1 was built for the scenario without RFID. It consists of two aisles each with 8 racks arranged in a matrix form as shown below. R represents Red and B represents blue types of entities.

The input distribution for the red entities is a “constant” arrival rate of 30 entities per arrival. ‘Days’ is chosen to be the units. Blue entities follow the same arrival pattern except that there are 35 entities per arrival. A sample screen shot for the Red entities is shown below. The process blocks titled “Rack 1 Aisle 1” have a Uniform (5, 7) distribution with units in hours. It is ensured that there is no queue in front of the racks to portray a real warehouse scenario.

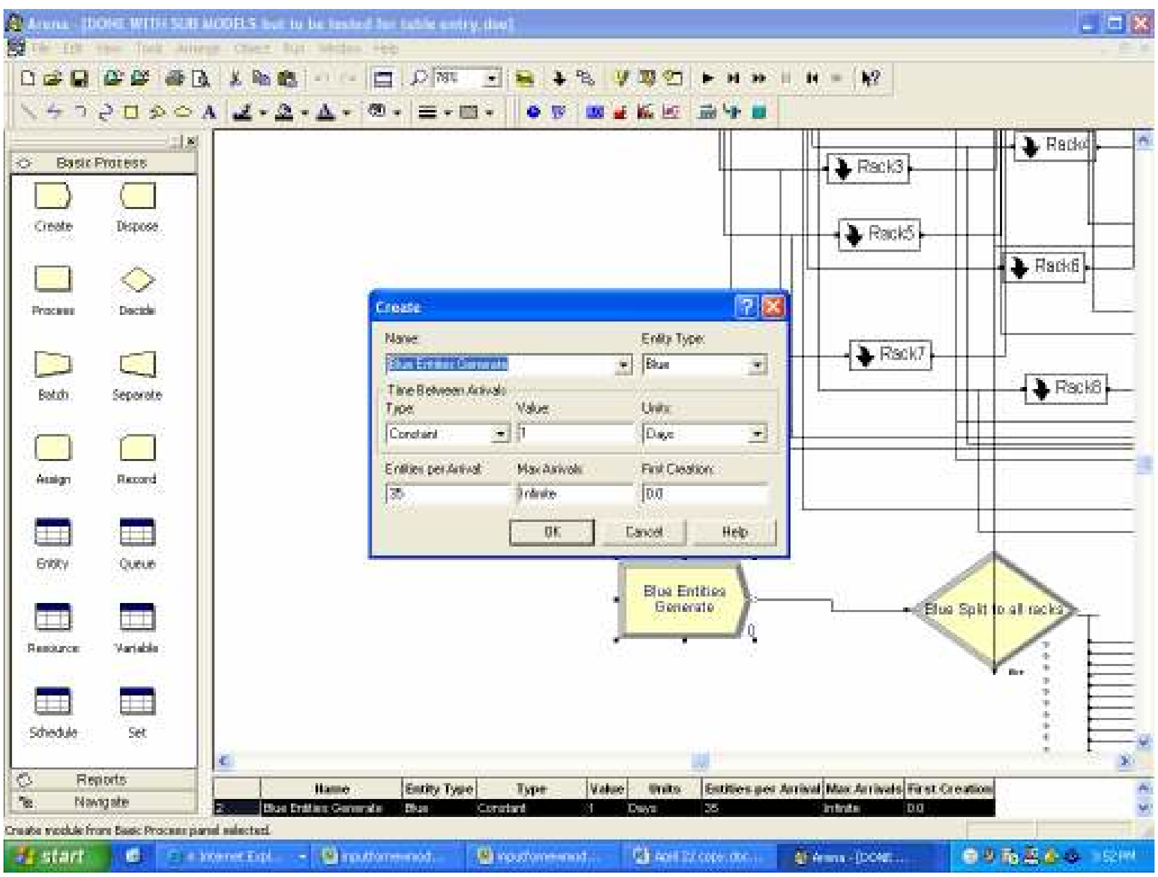

Figure A1a: Screen shot of ARENA model

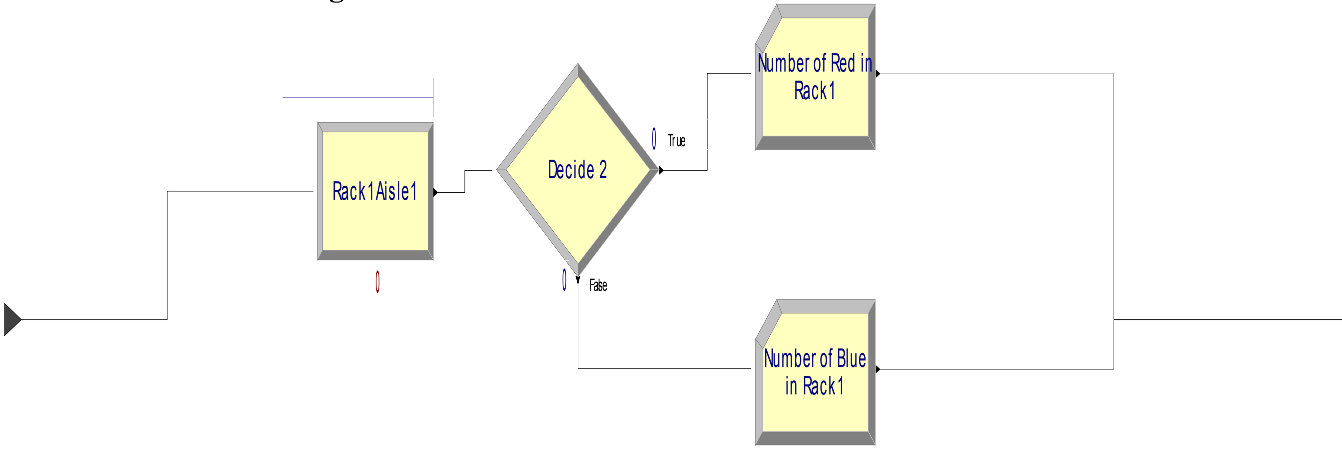

Figure A1b: Sub modules under each process block

R	R
R	R
B	**B**
B	**B**

Figure A1c: Aisle 1 with red and blue entities

The top four blocks are the storage region for the red and the bottom four is the storage region for the blue in Figure A1. The top four blocks are the storage region for the blue and the bottom four is the storage region for the red in Figure A2.

B	B
B	B
R	**R**
R	**R**

Figure A2: Aisle 2 with similar configurations

A product enters the warehouse and if there is space for it, is stored, else it is disposed. By running the model for 10 replications from t = 1 to t = 10 and clearing the statistics for every replication we ensure that the racks have 0 items in them before starting the next run.

Accuracy can be defined as the total percentage of Red (or Blue) items that are in their respective assigned positions *and* are detected by the RFID reader on scanning. The level of misplaced items remains constant. It is only the signal received from the tag that determines the read rate which in turn helps detect how many of each type has been read. It can be defined mathematically as the ratio of the Red items (or Blue) that was scanned and detected in racks assigned to Red (or Blue) to the total number of Red items (or Blue) in the warehouse (including misplaced Red items) at any given instance of time. Mathematically

$$\% \text{ Accuracy} = \frac{\text{Red items (or blue) in racks assigned to Red (or blue)}}{\text{Total number of Red items (or blue) that arrived for the simulation time}} \times 100$$

As the read rate increases the accuracy should also increase linearly. The following tables (Table A1 to Table A4) compare the scenarios for misplaced items when there is no RFID (i.e. a warehouse that doesn't have any readers and hence no data is collected on its misplaced or lost items) and when RFID readers with varying read rates are installed. The system without RFID corresponds to an actual scenario when products are not placed in their assigned locations (assigned by the WMS) or when they are moved internally and their location information is not updated. Therefore there is a decrease in the accuracy level (defined above) and hence searching is required.

However, if RFID readers are installed on the forklifts, the information about the number of items being moved from/to what storage location, can be obtained and updated in the WMS. This makes order picking operations extremely efficient and minimizes the expected travel/picking time. However, since the current RFID systems do have inherent problems the read rates of the RFID readers vary. Hence the *exact* number of items being moved or shipped may not be 100% accurate. To include that uncertainty and analyze what maybe the required *minimum* read rate to detect the *exact* number of misplaced items, we have 3 different read rate scenarios of 95%, 97% and 99%.

A.1.2 Results from system without RFID

Table A1 compares the % accuracy of red and blue items in the case where there is no RFID system installed.

Table A7: Comparison of accuracy for red and blue items without RFID

Simulation run number	**Red count in racks assigned to Red**	**Blue count in racks assigned to Blue**	**% Accuracy of Red**	**% Accuracy of Blue**
1	19	27	63.33	77.14
2	26	22	86.66	62.85
3	20	20	66.66	57.14
4	23	21	76.66	60
5	24	26	80	74.28
6	35	40	72.91	80

Simulation run number	Red count in racks assigned to Red	Blue count in racks assigned to Blue	% Accuracy of Red	% Accuracy of Blue
7	26	38	61.90	69.09
8	19	22	63.33	62.85
9	25	23	83.33	65.71
10	23	23	76.66	65.71

A.1.3 Results with 95% read rate - RFID readers installed

Table A2 summarizes the scenario in which the RFID reader has a 95% read rate i.e. the item is detected 95% of the time it is scanned.

Table A8: Comparison of accuracy for red and blue items with 95% read rate

Simulation run number	Red count in racks assigned to Red	Blue count in racks assigned to Blue	% Accuracy of Red	% Accuracy of Blue
1	16	27	53.33	77.14
2	24	20	80	57.14
3	19	16	63.33	45.71
4	21	21	70	60
5	22	26	73.33	74.28
6	35	38	72.91	76

Simulation run number	Red count in racks assigned to Red	Blue count in racks assigned to Blue	% Accuracy of Red	% Accuracy of Blue
7	25	34	59.52	61.81
8	18	21	60	60
9	23	20	76.66	57.14
10	20	22	66.66	62.85

Figure A3 depicts the comparison of RFID based system (95% read rate) and one without for Blue entities. As we can clearly see from the graph there exists differences in values between the red bars and blue dots. A system with 95% read rate is unable to detect the exact number of misplaced items and hence lags well behind the actual scenario. This is as good as not having the RFID system itself.

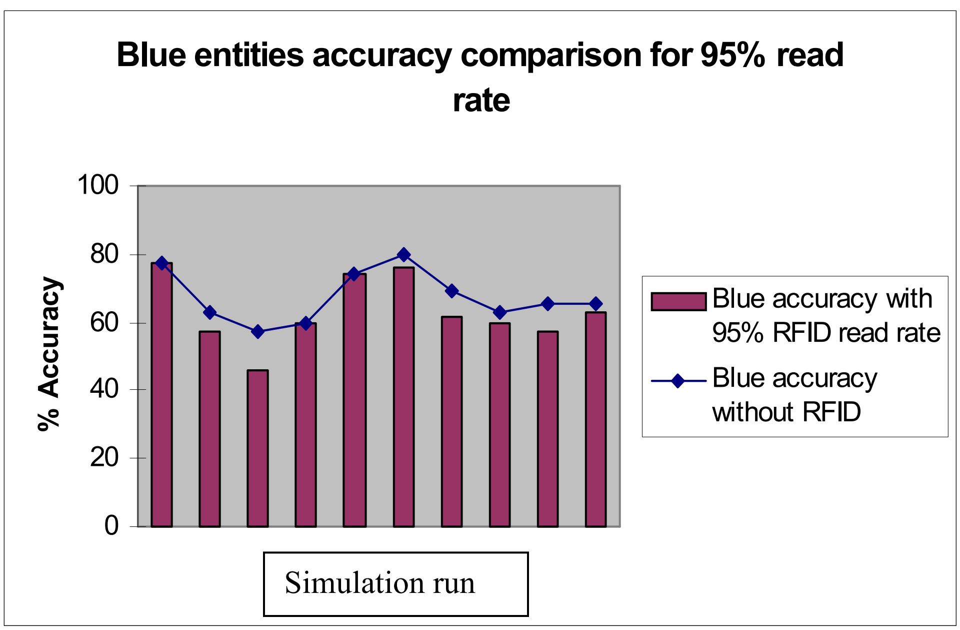

Figure A3: Comparison of accuracy of blue items for 95% read rate

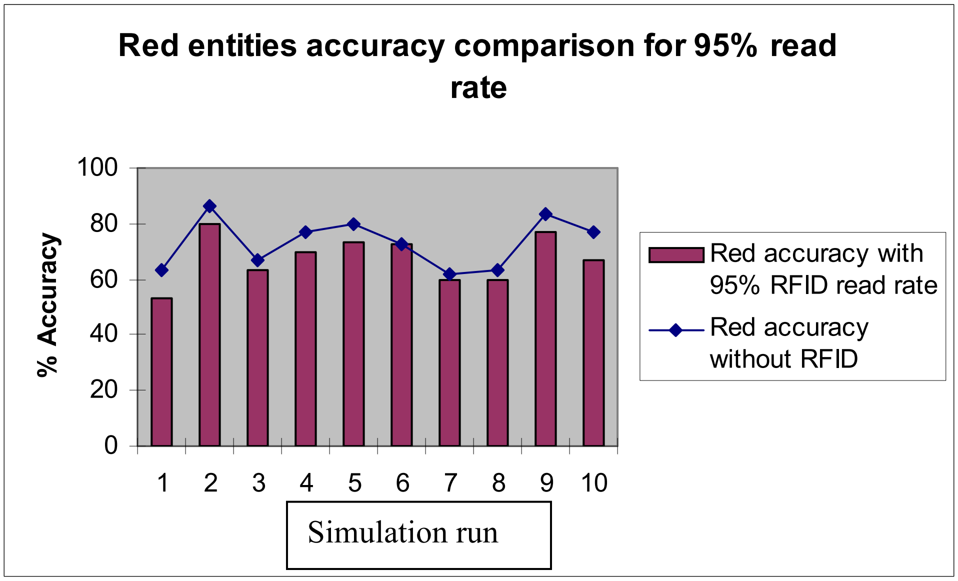

Figure A4: Comparison of accuracy of red items for 95% read rate

A.1.4 Results with 97% read rate - RFID readers installed

Table A3 summarizes the scenario in which the read rate of the readers has been increased to 97%. Figure A5 depicts the comparison of RFID based system (97% read rate) and one without for Blue entities. Compared to Figure A3, 97% read rate RFID system performs better but is still unable to detect all the misplaced items in the warehouse. There are many gaps (difference between the red bars and blue data points) between the actual values and the ones with RFID. Similarly, Figure A6 summarizes the variation of RFID based system (97% read rate) and one without for Blue entities.

TableA9: Comparison of accuracy for red and blue items with 97% read rate

Simulation run number	Red count in racks assigned to Red	Blue count in racks assigned to Blue	% Accuracy of Red	% Accuracy of Blue
1	18	27	60	77.14
2	24	21	80	60
3	19	19	63.33	54.28
4	23	19	76.66	54.28
5	24	25	80	71.42
6	35	38	72.91	76
7	26	37	61.90	67.27
8	18	22	60	62.85
9	25	22	83.33	62.85
10	22	22	73.33	62.85

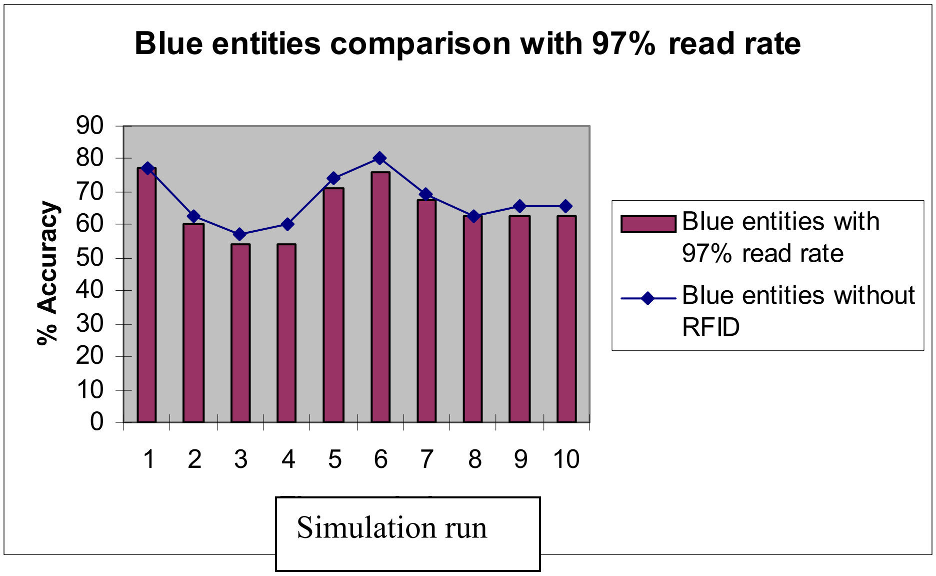

Figure A5: Comparison of accuracy of blue items for 97% read rate

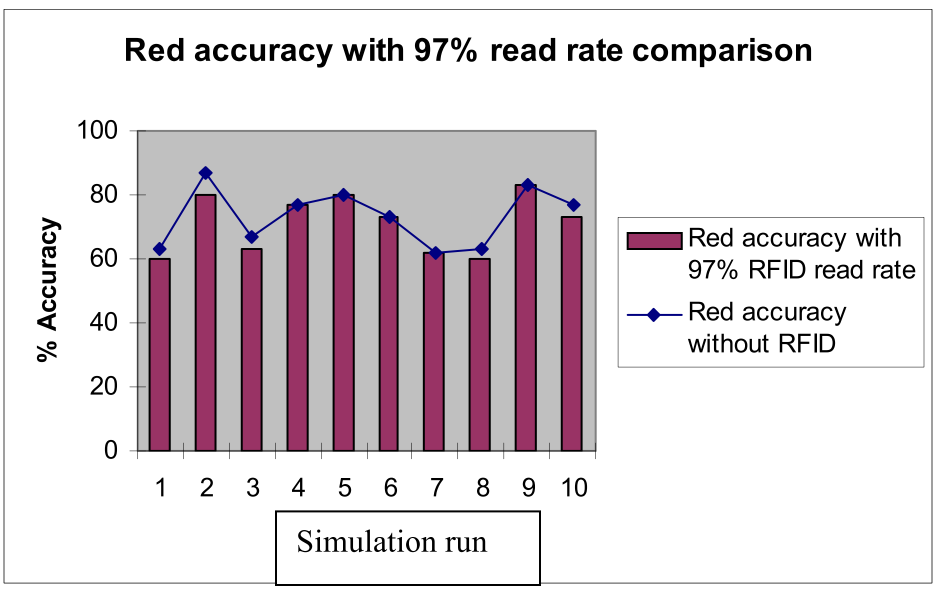

Figure A6: Comparison of accuracy of red items for 97% read rate

A.1.5 Results with 99% read rate - RFID readers installed

Table A4 summarizes the scenario in which the read rate of the readers has been assumed to be 99%. The following graph depicts the comparison of RFID based system (99% read rate) and one without for Blue entities. Increasing the read rate to 99% shows a drastic improvement in the readability and there are no value differences between the two cases. RFID is able to almost accurately capture the exact scenario of the misplaced items. Figures A7 and A8 show the variation for both entity types when 99% is the specified read rate.

Table A10: Comparison of accuracy for red and blue items with 99% read rate

Simulation run number	Red count in racks assigned to Red	Blue count in racks assigned to Blue	% Accuracy of Red	% Accuracy of Blue
1	19	26	63.33	77.14
2	26	22	86.66	62.85
3	20	20	66.66	57.14
4	23	21	76.66	60
5	24	26	80	74.28
6	35	40	72.91	80
7	26	37	50	67.27
8	19	22	63.33	62.85
9	25	23	83.33	65.71
10	22	23	73.33	65.71

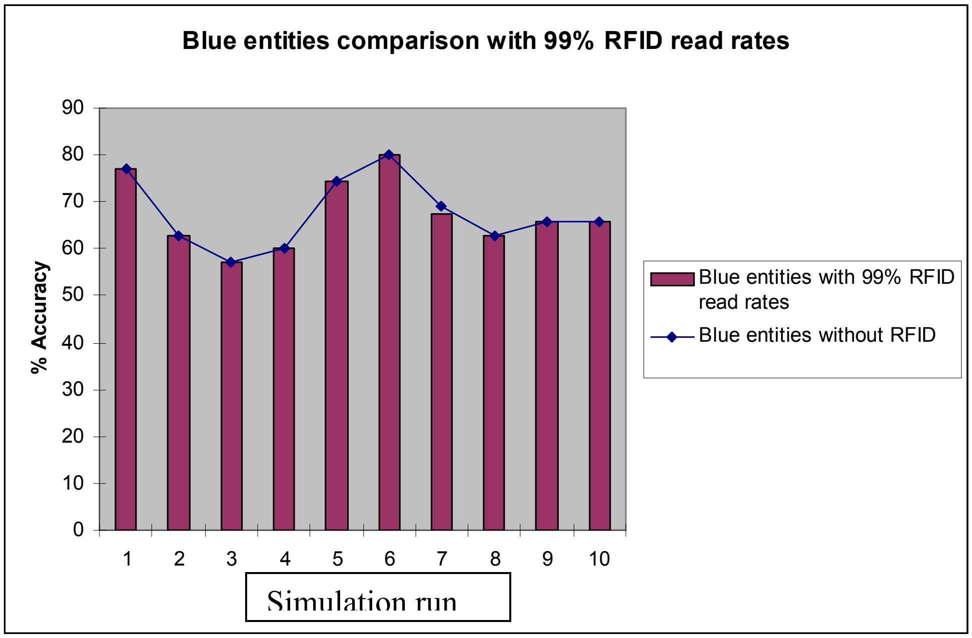

Figure A720: Comparison of accuracy of blue items for 99% read rate

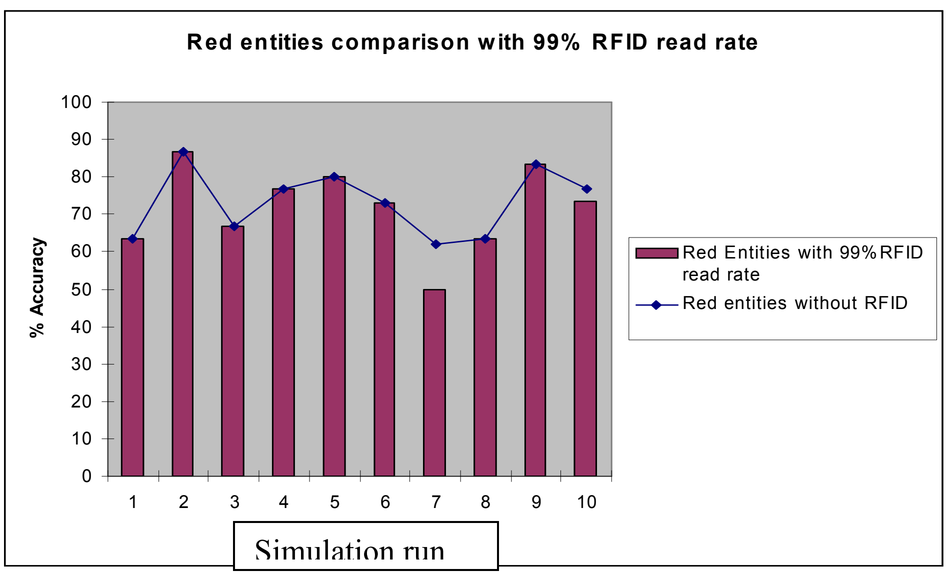

Figure 21: Comparison of accuracy of red items for 99% read rate

A.1.6 Explanation of results

From the above tables and graphs we can see that as the read rate of the RFID readers increase from 95% to 97%, the performance is much better. This could be clearly seen as the blue curve edges close to the red bars in the 97% read rate cases. This holds true for Red and Blue entities whose arrival and departure pattern are independent from each other in this simulation. Hence with a 97% read rate capability RFID readers could perform well for products that have no metal or liquid in them. As the read rate increases to 99%, in almost all the time periods, the values of entries without RFID matches the one with RFID. An RFID system with ~ 99 % read rate (readability) must be required to exactly capture the accurate number of misplaced items.

POMDP INPUT FILES:

POMDP model input parameters

Input Specification format file

discount: %f

values: [reward, cost]

states: [%d, <list of states>]

actions: [%d, <list of actions>]

observations: [%d, <list of observations>]

A SUBSET OF THE STATE TRANSITION FUNCTION

	T:up	0	1	2	3	4	5	6	7
0		0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
1		0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
2		0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
3		0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
4		0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
5		0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
6		0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
7		0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
8		0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0

9	**0.0**	**0.0**	**0.0**	**0.0**	**0.0**	**0.0**	**0.0**	**0.0**
10	**0.0**	**0.0**	**0.0**	**0.0**	**0.0**	**0.0**	**0.0**	**0.0**
11	**0.0**	**0.0**	**0.0**	**0.0**	**0.0**	**0.0**	**0.0**	**0.0**
12	**0.0**	**0.0**	**0.0**	**0.0**	**0.0**	**0.0**	**0.0**	**0.0**
13	**0.0**	**0.0**	**0.0**	**0.0**	**0.0**	**0.0**	**0.0**	**0.0**
14	**0.0**	**0.0**	**0.0**	**0.0**	**0.0**	**0.0**	**0.0**	**0.0**
15	**0.0**	**0.0**	**0.0**	**0.0**	**0.0**	**0.0**	**0.0**	**0.0**
16	**0.0**	**0.0**	**0.0**	**0.0**	**0.0**	**0.0**	**0.0**	**0.0**
17	**0.0**	**0.0**	**0.0**	**0.0**	**0.0**	**0.0**	**0.0**	**0.0**
18	**0.0**	**0.0**	**0.0**	**0.0**	**0.0**	**0.0**	**0.0**	**0.0**
19	**0.0**	**0.0**	**0.0**	**0.0**	**0.0**	**0.0**	**0.0**	**0.0**
20	**0.0**	**0.0**	**0.0**	**0.0**	**0.0**	**0.0**	**0.0**	**0.0**
21	**0.0**	**0.0**	**0.0**	**0.0**	**0.0**	**0.0**	**0.0**	**0.0**
22	**0.0**	**0.0**	**0.0**	**0.0**	**0.0**	**0.0**	**0.0**	**0.0**
23	**0.0**	**0.0**	**0.0**	**0.0**	**0.0**	**0.0**	**0.0**	**0.0**
24	**0.0**	**0.0**	**0.0**	**0.0**	**0.0**	**0.0**	**0.0**	**0.0**

OBSERVATION FUNCTION FOR MODEL GS1

O:0

0.0	0.0	0.0	1.0	0.0
0.0	0.0	0.0	1.0	0.0
0.0	0.0	0.0	1.0	0.0
0.0	0.0	0.0	1.0	0.0
0.0	0.0	0.0	1.0	0.0
0.0	0.0	0.0	1.0	0.0
0.0	0.0	0.0	1.0	0.0
0.0	0.0	0.0	1.0	0.0
0.0	0.0	0.0	1.0	0.0

0.0	0.0	0.0	1.0	0.0
0.0	0.0	0.0	1.0	0.0
0.0	0.0	0.0	1.0	0.0
0.0	0.0	0.0	1.0	0.0
0.0	0.0	0.0	1.0	0.0
0.0	0.0	0.0	1.0	0.0
0.0	0.0	0.0	1.0	0.0
0.0	0.0	0.0	1.0	0.0
0.0	0.0	0.0	1.0	0.0
0.0	0.0	0.0	1.0	0.0
0.0	0.0	0.0	1.0	0.0
0.0	0.0	0.0	1.0	0.0
0.0	0.0	0.8	0.2	0.0
0.0	0.1	0.8	0.1	0.0
0.0	0.0	0.8	0.2	0.0
1.0	0.0	0.0	0.0	0.0
0.0	0.8	0.2	0.0	0.0
0.0	0.8	0.1	0.1	0.0
0.0	0.0	0.0	1.0	0.0
0.0	0.1	0.8	0.1	0.0
0.0	0.0	0.8	0.2	0.0
0.0	0.1	0.8	0.1	0.0
0.0	0.8	0.2	0.0	0.0
1.0	0.0	0.0	0.0	0.0
0.0	0.8	0.2	0.0	0.0
0.0	0.0	0.0	1.0	0.0
0.0	0.0	0.8	0.2	0.0
0.0	0.1	0.8	0.1	0.0
0.0	0.0	0.8	0.2	0.0
0.0	0.8	0.1	0.1	0.0
0.0	0.8	0.2	0.0	0.0
1.0	0.0	0.0	0.0	0.0
0.0	0.0	0.0	1.0	0.0
0.0	0.0	0.0	1.0	0.0
0.0	0.0	0.0	0.0	1.0

O:1

1.0	0.0	0.0	0.0	0.0
0.0	0.8	0.2	0.0	0.0
0.0	0.8	0.1	0.1	0.0
0.0	0.0	0.8	0.2	0.0
0.0	0.1	0.8	0.1	0.0
0.0	0.0	0.8	0.2	0.0
0.0	0.0	0.0	1.0	0.0
0.0	0.8	0.2	0.0	0.0
1.0	0.0	0.0	0.0	0.0
0.0	0.8	0.2	0.0	0.0
0.0	0.1	0.8	0.1	0.0

0.0	0.0	0.8	0.2	0.0
0.0	0.1	0.8	0.1	0.0
0.0	0.0	0.0	1.0	0.0
0.0	0.8	0.1	0.1	0.0
0.0	0.8	0.2	0.0	0.0
1.0	0.0	0.0	0.0	0.0
0.0	0.0	0.8	0.2	0.0
0.0	0.1	0.8	0.1	0.0
0.0	0.0	0.8	0.2	0.0
0.0	0.0	0.0	1.0	0.0
0.0	0.0	0.0	1.0	0.0
0.0	0.0	0.0	1.0	0.0
0.0	0.0	0.0	1.0	0.0
0.0	0.0	0.0	1.0	0.0
0.0	0.0	0.0	1.0	0.0
0.0	0.0	0.0	1.0	0.0
0.0	0.0	0.0	1.0	0.0
0.0	0.0	0.0	1.0	0.0
0.0	0.0	0.0	1.0	0.0
0.0	0.0	0.0	1.0	0.0
0.0	0.0	0.0	1.0	0.0
0.0	0.0	0.0	1.0	0.0
0.0	0.0	0.0	1.0	0.0
0.0	0.0	0.0	1.0	0.0
0.0	0.0	0.0	1.0	0.0
0.0	0.0	0.0	1.0	0.0
0.0	0.0	0.0	1.0	0.0
0.0	0.0	0.0	1.0	0.0
0.0	0.0	0.0	1.0	0.0
0.0	0.0	0.0	1.0	0.0
0.0	0.0	0.0	1.0	0.0
0.0	0.0	0.0	1.0	0.0
0.0	0.0	0.0	0.0	1.0

O:2

1.0	0.0	0.0	0.0	0.0
0.0	0.1	0.9	0.0	0.0
0.0	0.1	0.8	0.1	0.0
0.0	0.8	0.1	0.1	0.0
0.0	0.1	0.8	0.1	0.0
0.0	0.0	0.8	0.2	0.0
0.0	0.0	0.0	1.0	0.0
0.0	0.8	0.2	0.0	0.0
1.0	0.0	0.0	0.0	0.0
0.0	0.1	0.9	0.0	0.0
0.0	0.7	0.15	0.15	0.0
0.0	0.8	0.1	0.1	0.0
0.0	0.1	0.8	0.1	0.0
0.0	0.0	0.0	1.0	0.0

0.0	0.0	0.0	1.0	0.0
0.0	0.0	0.0	1.0	0.0
0.0	0.0	0.0	1.0	0.0
0.0	0.0	0.0	1.0	0.0
0.0	0.0	0.0	1.0	0.0
0.0	0.0	0.0	1.0	0.0
0.0	0.0	0.0	1.0	0.0
0.0	0.8	0.1	0.1	0.0
0.0	0.1	0.8	0.1	0.0
0.0	0.0	0.8	0.2	0.0
1.0	0.0	0.0	0.0	0.0
0.0	0.1	0.9	0.0	0.0
0.0	0.0	0.8	0.2	0.0
0.0	0.0	0.0	1.0	0.0
0.0	0.7	0.15	0.15	0.0
0.0	0.8	0.1	0.1	0.0
0.0	0.1	0.8	0.1	0.0
0.0	0.8	0.2	0.0	0.0
1.0	0.0	0.0	0.0	0.0
0.0	0.1	0.9	0.0	0.0
0.0	0.0	0.0	1.0	0.0
0.0	0.0	0.0	1.0	0.0
0.0	0.0	0.0	1.0	0.0
0.0	0.0	0.0	1.0	0.0
0.0	0.0	0.0	1.0	0.0
0.0	0.0	0.0	1.0	0.0
0.0	0.0	0.0	1.0	0.0
0.0	0.0	0.0	1.0	0.0
0.0	0.0	0.0	1.0	0.0
0.0	0.0	0.0	0.0	1.0

O:3

0.0	0.0	0.0	1.0	0.0
0.0	0.0	0.0	1.0	0.0
0.0	0.0	0.0	1.0	0.0
0.0	0.0	0.0	1.0	0.0
0.0	0.0	0.0	1.0	0.0
0.0	0.0	0.0	1.0	0.0
0.0	0.0	0.0	1.0	0.0
0.0	0.1	0.9	0.0	0.0
1.0	0.0	0.0	0.0	0.0
0.0	0.8	0.2	0.0	0.0
0.0	0.1	0.8	0.1	0.0
0.0	0.8	0.1	0.1	0.0
0.0	0.7	0.15	0.15	0.0
0.0	0.0	0.0	1.0	0.0
0.0	0.1	0.8	0.1	0.0
0.0	0.1	0.9	0.0	0.0

1.0	0.0	0.0	0.0	0.0
0.0	0.0	0.8	0.2	0.0
0.0	0.1	0.8	0.1	0.0
0.0	0.8	0.1	0.1	0.0
0.0	0.0	0.0	1.0	0.0
0.0	0.0	0.0	1.0	0.0
0.0	0.0	0.0	1.0	0.0
0.0	0.0	0.0	1.0	0.0
0.0	0.0	0.0	1.0	0.0
0.0	0.0	0.0	1.0	0.0
0.0	0.0	0.0	1.0	0.0
0.0	0.0	0.0	1.0	0.0
0.0	0.1	0.8	0.1	0.0
0.0	0.8	0.1	0.1	0.0
0.0	0.7	0.15	0.15	0.0
0.0	0.1	0.9	0.0	0.0
1.0	0.0	0.00	0.00	0.0
0.0	0.8	0.2	0.0	0.0
0.0	0.0	0.0	1.0	0.0
0.0	0.0	0.8	0.2	0.0
0.0	0.1	0.8	0.1	0.0
0.0	0.8	0.1	0.1	0.0
0.0	0.1	0.8	0.1	0.0
0.0	0.1	0.9	0.0	0.0
1.0	0.0	0.0	0.0	0.0
0.0	0.0	0.0	1.0	0.0
0.0	0.0	0.0	1.0	0.0
0.0	0.0	0.0	0.0	1.0

OBSERVATION FUNCTION FOR MODEL GS2 (LESSER UNCERTAINTY)

O:0:0

0.0	0.0	0.0	1.0	0.0

O:0:1

0.0	0.0	0.0	1.0	0.0

O:0:2

0.0	0.0	0.0	1.0	0.0

O:0:3

0.0	0.0	0.0	1.0	0.0

O:0:4

0.0	0.0	0.0	1.0	0.0

O:0:5

0.0	0.0	0.0	1.0	0.0

O:0:6

0.0	0.0	0.0	1.0	0.0

O:0:7

0.0	0.0	0.0	1.0	0.0

O:0:8

0.0	0.0	0.0	1.0	0.0

O:0:9

0.0	0.0	0.0	1.0	0.0

O:0:10

0.0	0.0	0.0	1.0	0.0

O:0:11

0.0	0.0	0.0	1.0	0.0

O:0:12

0.0	0.0	0.0	1.0	0.0

O:0:13

0.0	0.0	0.0	1.0	0.0

O:0:14

0.0	0.0	0.0	1.0	0.0

O:0:15

0.0	0.0	0.0	1.0	0.0

O:0:16

0.0	0.0	0.0	1.0	0.0

O:0:17

0.0	0.0	0.0	1.0	0.0

O:0:18

0.0	0.0	0.0	1.0	0.0

O:0:19

0.0	0.0	0.0	1.0	0.0

O:0:20

0.0	0.0	0.0	1.0	0.0

O:0:21

0.0	0.1	0.9	0.0	0.0

O:0:22

0.0	0.3	0.7	0.0	0.0

O:0:23

0.0	0.2	0.8	0.0	0.0

O:0:24

1.0	0.0	0.0	0.0	0.0

O:0:25

0.0	0.8	0.2	0.0	0.0

O:0:26

0.0	0.7	0.3	0.0	0.0

O:0:27

0.0	0.0	0.0	1.0	0.0

O:0:28

0.0	0.3	0.7	0.0	0.0

O:0:29

0.0	0.1	0.9	0.0	0.0

O:0:30

0.0	0.3	0.7	0.0	0.0

O:0:31	0.0	0.8	0.2	0.0	0.0
O:0:32	1.0	0.0	0.0	0.0	0.0
O:0:33	0.0	0.8	0.2	0.0	0.0
O:0:34	0.0	0.0	0.0	1.0	0.0
O:0:35	0.0	0.2	0.8	0.0	0.0
O:0:36	0.0	0.3	0.7	0.0	0.0
O:0:37	0.0	0.1	0.9	0.0	0.0
O:0:38	0.0	0.7	0.3	0.0	0.0
O:0:39	0.0	0.8	0.2	0.0	0.0
O:0:40	1.0	0.0	0.0	0.0	0.0
O:0:41	0.0	0.0	0.0	1.0	0.0
O:0:42	0.0	0.0	0.0	1.0	0.0

O:1:0	1.0	0.0	0.0	0.0	0.0
O:1:1	0.0	0.8	0.2	0.0	0.0
O:1:2	0.0	0.7	0.3	0.0	0.0
O:1:3	0.0	0.1	0.9	0.0	0.0
O:1:4	0.0	0.3	0.7	0.0	0.0
O:1:5	0.0	0.2	0.8	0.0	0.0
O:1:6	0.0	0.0	0.0	1.0	0.0
O:1:7	0.0	0.8	0.2	0.0	0.0
O:1:8	1.0	0.0	0.0	0.0	0.0
O:1:9	0.0	0.8	0.2	0.0	0.0
O:1:10	0.0	0.3	0.7	0.0	0.0
O:1:11	0.0	0.1	0.9	0.0	0.0
O:1:12					

	0.0	0.3	0.7	0.0	0.0
O:1:13					
	0.0	0.0	0.0	1.0	0.0
O:1:14					
	0.0	0.7	0.3	0.0	0.0
O:1:15					
	0.0	0.8	0.2	0.0	0.0
O:1:16					
	1.0	0.0	0.0	0.0	0.0
O:1:17					
	0.0	0.2	0.8	0.0	0.0
O:1:18					
	0.0	0.3	0.8	0.0	0.0
O:1:19					
	0.0	0.1	0.9	0.0	0.0
O:1:20					
	0.0	0.0	0.0	1.0	0.0
O:1:21					
	0.0	0.0	0.0	1.0	0.0
O:1:22					
	0.0	0.0	0.0	1.0	0.0
O:1:23					
	0.0	0.0	0.0	1.0	0.0
O:1:24					
	0.0	0.0	0.0	1.0	0.0
O:1:25					
	0.0	0.0	0.0	1.0	0.0
O:1:26					
	0.0	0.0	0.0	1.0	0.0
O:1:27					
	0.0	0.0	0.0	1.0	0.0
O:1:28					
	0.0	0.0	0.0	1.0	0.0
O:1:29					
	0.0	0.0	0.0	1.0	0.0
O:1:30					
	0.0	0.0	0.0	1.0	0.0
O:1:31					
	0.0	0.0	0.0	1.0	0.0
O:1:32					
	0.0	0.0	0.0	1.0	0.0
O:1:33					
	0.0	0.0	0.0	1.0	0.0
O:1:34					
	0.0	0.0	0.0	1.0	0.0
O:1:35					
	0.0	0.0	0.0	1.0	0.0
O:1:36					
	0.0	0.0	0.0	1.0	0.0

O:1:37
0.0 0.0 0.0 1.0 0.0
O:1:38
0.0 0.0 0.0 1.0 0.0
O:1:39
0.0 0.0 0.0 1.0 0.0
O:1:40
0.0 0.0 0.0 1.0 0.0
O:1:41
0.0 0.0 0.0 1.0 0.0
O:1:42
0.0 0.0 0.0 1.0 0.0

O:2:0
1.0 0.0 0.0 0.0 0.0
O:2:1
0.0 0.1 0.9 0.0 0.0
O:2:2
0.0 0.3 0.7 0.0 0.0
O:2:3
0.0 0.7 0.3 0.0 0.0
O:2:4
0.0 0.3 0.7 0.0 0.0
O:2:5
0.0 0.2 0.8 0.0 0.0
O:2:6
0.0 0.0 0.0 1.0 0.0
O:2:7
0.0 0.8 0.2 0.0 0.0
O:2:8
1.0 0.0 0.0 0.0 0.0
O:2:9
0.0 0.1 0.9 0.0 0.0
O:2:10
0.0 0.7 0.3 0.0 0.0
O:2:11
0.0 0.8 0.2 0.0 0.0
O:2:12
0.0 0.3 0.7 0.0 0.0
O:2:13
0.0 0.0 0.0 1.0 0.0
O:2:21
0.0 0.8 0.2 0.0 0.0
O:2:22
0.0 0.3 0.7 0.0 0.0
O:2:23
0.0 0.2 0.8 0.0 0.0

O:2:24
1.0 0.0 0.0 0.0 0.0
O:2:25
0.0 0.1 0.9 0.0 0.0
O:2:26
0.0 0.3 0.7 0.0 0.0
O:2:27
0.0 0.0 0.0 1.0 0.0
O:2:28
0.0 0.7 0.3 0.0 0.0
O:2:29
0.0 0.8 0.2 0.0 0.0
O:2:30
0.0 0.2 0.8 0.0 0.0
O:2:31
0.0 0.8 0.2 0.0 0.0
O:2:32
1.0 0.0 0.0 0.0 0.0
O:2:33
0.0 0.1 0.9 0.0 0.0
O:2:34
0.0 0.0 0.0 1.0 0.0
O:2:14
0.0 0.0 0.0 1.0 0.0
O:2:15
0.0 0.0 0.0 1.0 0.0
O:2:16
0.0 0.0 0.0 1.0 0.0
O:2:17
0.0 0.0 0.0 1.0 0.0
O:2:18
0.0 0.0 0.0 1.0 0.0
O:2:19
0.0 0.0 0.0 1.0 0.0
O:2:20
0.0 0.0 0.0 1.0 0.0
O:2:35
0.0 0.0 0.0 1.0 0.0
O:2:36
0.0 0.0 0.0 1.0 0.0
O:2:37
0.0 0.0 0.0 1.0 0.0
O:2:38
0.0 0.0 0.0 1.0 0.0
O:2:39
0.0 0.0 0.0 1.0 0.0
O:2:40
0.0 0.0 0.0 1.0 0.0
O:2:41
0.0 0.0 0.0 1.0 0.0

O:2:42

0.0	0.0	0.0	1.0	0.0

O:3:0

0.0	0.0	0.0	1.0	0.0

O:3:1

0.0	0.0	0.0	1.0	0.0

O:3:2

0.0	0.0	0.0	1.0	0.0

O:3:3

0.0	0.0	0.0	1.0	0.0

O:3:4

0.0	0.0	0.0	1.0	0.0

O:3:5

0.0	0.0	0.0	1.0	0.0

O:3:6

0.0	0.0	0.0	1.0	0.0

O:3:21

0.0	0.0	0.0	1.0	0.0

O:3:22

0.0	0.0	0.0	1.0	0.0

O:3:23

0.0	0.0	0.0	1.0	0.0

O:3:24

0.0	0.0	0.0	1.0	0.0

O:3:25

0.0	0.0	0.0	1.0	0.0

O:3:26

0.0	0.0	0.0	1.0	0.0

O:3:27

0.0	0.0	0.0	1.0	0.0

O:3:7

0.0	0.1	0.9	0.0	0.0

O:3:8

1.0	0.0	0.0	0.0	0.0

O:3:9

0.0	0.8	0.2	0.0	0.0

O:3:10

0.0	0.2	0.8	0.0	0.0

O:3:11

0.0	0.8	0.2	0.0	0.0

O:3:12

0.0	0.8	0.2	0.0	0.0

O:3:13

0.0	0.0	0.0	1.0	0.0

O:3:14

0.0	0.2	0.8	0.0	0.0

O:3:15

0.0	0.1	0.9	0.0	0.0

O:3:16

1.0	0.0	0.0	0.0	0.0

O:3:17

0.0	0.2	0.8	0.0	0.0

O:3:18

0.0	0.3	0.7	0.0	0.0

O:3:19

0.0	0.8	0.2	0.0	0.0

O:3:20

0.0	0.0	0.0	1.0	0.0

O:3:28

0.0	0.2	0.8	0.0	0.0

O:3:29

0.0	0.8	0.2	0.0	0.0

O:3:30

0.0	0.7	0.3	0.0	0.0

O:3:31

0.0	0.1	0.9	0.0	0.0

O:3:32

1.0	0.0	0.0	0.0	0.0

O:3:33

0.0	0.8	0.2	0.0	0.0

O:3:34

0.0	0.0	0.0	1.0	0.0

O:3:35

0.0	0.0	1.0	0.0	0.0

O:3:36

0.0	0.0	1.0	0.0	0.0

O:3:37

0.0	1.0	0.0	0.0	0.0

O:3:38

0.0	0.0	1.0	0.0	0.0

O:3:39

0.0	0.0	1.0	0.0	0.0

O:3:40

1.0	0.0	0.0	0.0	0.0

O:3:41

0.0	0.0	0.0	1.0	0.0

O:3:42

0.0	0.0	0.0	1.0	0.0

OBSERVATION FUNCTION FOR POOR OBSERVATION MODEL

```
          O:0:0
0.0   0.0   0.0   1.0 0.0
          O:0:1
0.0   0.0   0.0   1.0 0.0
```

```
        O:0:2
0.0   0.0   0.0   1.0 0.0
        O:0:3
0.0   0.0   0.0   1.0 0.0
        O:0:4
0.0   0.0   0.0   1.0 0.0
        O:0:5
0.0   0.0   0.0   1.0 0.0
        O:0:6
0.0   0.0   0.0   1.0 0.0
        O:0:7
0.0   0.0   0.0   1.0 0.0
        O:0:8
0.0   0.0   0.0   1.0 0.0
        O:0:9
0.0   0.0   0.0   1.0 0.0
       O:0:10
0.0   0.0   0.0   1.0 0.0
       O:0:11
0.0   0.0   0.0   1.0 0.0
       O:0:12
0.0   0.0   0.0   1.0 0.0
       O:0:13
0.0   0.0   0.0   1.0 0.0
       O:0:14
0.0   0.0   0.0   1.0 0.0
       O:0:15
0.0   0.0   0.0   1.0 0.0
       O:0:16
0.0   0.0   0.0   1.0 0.0
       O:0:17
0.0   0.0   0.0   1.0 0.0
       O:0:18
0.0   0.0   0.0   1.0 0.0
       O:0:19
0.0   0.0   0.0   1.0 0.0
       O:0:20
0.0   0.0   0.0   1.0 0.0

       O:0:21
0.0   0.0   0.5   0.5 0.0
       O:0:22
0.0   0.0   0.5   0.5 0.0
       O:0:23
0.0   0.0   0.5   0.5 0.0
       O:0:24
1.0   0.0   0.0   0.0 0.0
       O:0:25
0.0   0.5   0.5   0.0 0.0
       O:0:26
0.0   0.5   0.5   0.0 0.0
       O:0:27
0.0   0.0   0.0   1.0 0.0
       O:0:28
0.0   0.0   0.5   0.5 0.0
       O:0:29
0.0   0.0   0.5   0.5 0.0
```

O:0:30
0.0 0.0 0.5 0.5 0.0
O:0:31
0.0 0.5 0.5 0.0 0.0
O:0:32
1.0 0.0 0.0 0.0 0.0
O:0:33
0.0 0.5 0.5 0 0.0
O:0:34
0.0 0.0 0.0 1.0 0.0
O:0:35
0.0 0.0 0.5 0.5 0.0
O:0:36
0.0 0.0 0.5 0.5 0.0
O:0:37
0.0 0.0 0.5 0.5 0.0
O:0:38
0.0 0.5 0.5 0.0 0.0
O:0:39
0.0 0.5 0.5 0.0 0.0
O:0:40
1.0 0.0 0.0 0.0 0.0
O:0:41
0.0 0.0 0.0 1.0 0.0
O:0:42
0.0 0.0 0.0 1.0 0.0

O:1:0
1.0 0.0 0.0 0.0 0.0
O:1:1
0.0 0.5 0.5 0.0 0.0
O:1:2
0.0 0.5 0.5 0.0 0.0
O:1:3
0.0 0.0 0.5 0.5 0.0
O:1:4
0.0 0.0 0.5 0.5 0.0
O:1:5
0.0 0.0 0.5 0.5 0.0
O:1:6
0.0 0.0 0.0 1.0 0.0
O:1:7
0.0 0.5 0.5 0.0 0.0
O:1:8
1.0 0.0 0.0 0.0 0.0
O:1:9
0.0 0.5 0.5 0.0 0.0
O:1:10
0.0 0.0 0.5 0.5 0.0
O:1:11
0.0 0.0 0.5 0.5 0.0
O:1:12
0.0 0.0 0.5 0.5 0.0
O:1:13
0.0 0.0 0.0 1.0 0.0
O:1:14
0.0 0.5 0.5 0.0 0.0

```
          O:1:15
0.0    0.5   0.5    0.0 0.0
          O:1:16
1.0    0.0   0.0    0.0 0.0
          O:1:17
0.0    0.0   0.5    0.5 0.0
          O:1:18
0.0    0.0   0.5    0.5 0.0
          O:1:19
0.0    0.0   0.5    0.5 0.0
          O:1:20
0.0    0.0   0.0    1.0 0.0

          O:1:21
0.0    0.0   0.0    1.0 0.0
          O:1:22
0.0    0.0   0.0    1.0 0.0
          O:1:23
0.0    0.0   0.0    1.0 0.0
          O:1:24
0.0    0.0   0.0    1.0 0.0
          O:1:25
0.0    0.0   0.0    1.0 0.0
          O:1:26
0.0    0.0   0.0    1.0 0.0
          O:1:27
0.0    0.0   0.0    1.0 0.0
          O:1:28
0.0    0.0   0.0    1.0 0.0
          O:1:29
0.0    0.0   0.0    1.0 0.0
          O:1:30
0.0    0.0   0.0    1.0 0.0
          O:1:31
0.0    0.0   0.0    1.0 0.0
          O:1:32
0.0    0.0   0.0    1.0 0.0
          O:1:33
0.0    0.0   0.0    1.0 0.0
          O:1:34
0.0    0.0   0.0    1.0 0.0
          O:1:35
0.0    0.0   0.0    1.0 0.0
          O:1:36
0.0    0.0   0.0    1.0 0.0
          O:1:37
0.0    0.0   0.0    1.0 0.0
          O:1:38
0.0    0.0   0.0    1.0 0.0
          O:1:39
0.0    0.0   0.0    1.0 0.0
          O:1:40
0.0    0.0   0.0    1.0 0.0
          O:1:41
0.0    0.0   0.0    1.0 0.0
          O:1:42
0.0    0.0   0.0    1.0 0.0
```

```
            O:2:0
1.0    0.0    0.0    0.0 0.0
            O:2:1
0.0    0.0    1.0    0.0 0.0
            O:2:2
0.0    0.0    0.5    0.5 0.0
            O:2:3
0.0    0.5    0.5    0.0 0.0
            O:2:4
0.0    0.0    0.5    0.5 0.0
            O:2:5
0.0    0.0    0.5    0.5 0.0
            O:2:6
0.0    0.0    0.0    1.0 0.0
            O:2:7
0.0    0.5    0.5    0.0 0.0
            O:2:8
1.0    0.0    0.0    0.0 0.0
            O:2:9
0.0    0.0    0.5    0.5 0.0
           O:2:10
0.0    0.5    0.5    0.0 0.0
           O:2:11
0.0    0.5    0.5    0.0 0.0
           O:2:12
0.0    0.0    0.5    0.5 0.0
           O:2:13
0.0    0.0    0.0    1.0 0.0
           O:2:21
0.0    0.5    0.5    0.0 0.0
           O:2:22
0.0    0.0    0.5    0.5 0.0
           O:2:23
0.0    0.0    0.5    0.5 0.0
           O:2:24
1.0    0.0    0.0    0.0 0.0
           O:2:25
0.0    0.0    0.5    0.5 0.0
           O:2:26
0.0    0.0    0.5    0.5 0.0
           O:2:27
0.0    0.0    0.0    1.0 0.0
           O:2:28
0.0    0.5    0.5    0.0 0.0
           O:2:29
0.0    0.5    0.5    0.0 0.0
           O:2:30
0.0    0.0    0.5    0.5 0.0
           O:2:31
0.0    0.5    0.5    0.0 0.0
           O:2:32
1.0    0.0    0.0    0.0 0.0
           O:2:33
0.0    0.0    0.5    0.5 0.0
           O:2:34
0.0    0.0    0.0    1.0 0.0
           O:2:14
```

```
0.0    0.0    0.0    1.0 0.0
           O:2:15
0.0    0.0    0.0    1.0 0.0
           O:2:16
0.0    0.0    0.0    1.0 0.0
           O:2:17
0.0    0.0    0.0    1.0 0.0
           O:2:18
0.0    0.0    0.0    1.0 0.0
           O:2:19
0.0    0.0    0.0    1.0 0.0
           O:2:20
0.0    0.0    0.0    1.0 0.0
           O:2:35
0.0    0.0    0.0    1.0 0.0
           O:2:36
0.0    0.0    0.0    1.0 0.0
           O:2:37
0.0    0.0    0.0    1.0 0.0
           O:2:38
0.0    0.0    0.0    1.0 0.0
           O:2:39
0.0    0.0    0.0    1.0 0.0
           O:2:40
0.0    0.0    0.0    1.0 0.0
           O:2:41
0.0    0.0    0.0    1.0 0.0
           O:2:42
0.0    0.0    0.0    1.0 0.0
            O:3:0
0.0    0.0    0.0    1.0 0.0
            O:3:1
0.0    0.0    0.0    1.0 0.0
            O:3:2
0.0    0.0    0.0    1.0 0.0
            O:3:3
0.0    0.0    0.0    1.0 0.0
            O:3:4
0.0    0.0    0.0    1.0 0.0
            O:3:5
0.0    0.0    0.0    1.0 0.0
            O:3:6
0.0    0.0    0.0    1.0 0.0
           O:3:21
0.0    0.0    0.0    1.0 0.0
           O:3:22
0.0    0.0    0.0    1.0 0.0
           O:3:23
0.0    0.0    0.0    1.0 0.0
           O:3:24
0.0    0.0    0.0    1.0 0.0
           O:3:25
0.0    0.0    0.0    1.0 0.0
           O:3:26
0.0    0.0    0.0    1.0 0.0
           O:3:27
0.0    0.0    0.0    1.0 0.0
```

```
        O:3:7
0.0   0.0   0.5   0.5 0.0
        O:3:8
1.0   0.0   0.0   0.0 0.0
        O:3:9
0.0   0.5   0.5   0.0 0.0
       O:3:10
0.0   0.0   0.5   0.5 0.0
       O:3:11
0.0   0.5   0.5   0.0 0.0
       O:3:12
0.0   0.5   0.5   0.0 0.0
       O:3:13
0.0   0.0   0.0   1.0 0.0
       O:3:14
0.0   0.0   0.5   0.5 0.0
       O:3:15
0.0   0.0   0.5   0.5 0.0
       O:3:16
1.0   0.0   0.0   0.0 0.0
       O:3:17
0.0   0.0   0.5   0.5 0.0
       O:3:18
0.0   0.0   0.5   0.5 0.0
       O:3:19
0.0   0.5   0.5   0.0 0.0
       O:3:20
0.0   0.0   0.0   1.0 0.0
       O:3:28
0.0   0.0   0.5   0.5 0.0
       O:3:29
0.0   0.5   0.5   0.0 0.0
       O:3:30
0.0   0.5   0.5   0.0 0.0
       O:3:31
0.0   0.0   0.5   0.5 0.0
       O:3:32
1.0   0.0   0.0   0.0 0.0
       O:3:33
0.0   0.5   0.5   0.0 0.0
       O:3:34
0.0   0.0   0.0   1.0 0.0
       O:3:35
0.0   0.0   0.5   0.5 0.0
       O:3:36
0.0   0.0   0.5   0.5 0.0
       O:3:37
0.0   0.5   0.5   0.0 0.0
       O:3:38
0.0   0.0   0.5   0.5 0.0
       O:3:39
0.0   0.0   0.5   0.5 0.0
       O:3:40
1.0   0.0   0.0   0.0 0.0
       O:3:41
0.0   0.0   0.0   1.0 0.0
```

```
O:3:42
0.0    0.0    0.0    1.0 0.0
```

Reward model:

```
R:up:21:*:* -10.0
R:up:22:*:* -10.0
R:up:23:*:* -10.0
R:up:24:*:* -10.0
R:up:25:*:* -10.0
R:up:26:*:* -10.0
R:up:27:*:* -10.0
R:up:28:*:* -10.0
R:up:29:*:* -10.0
R:up:30:*:* -10.0
R:up:31:*:* -10.0
R:up:32:*:* -10.0
R:up:33:*:* -10.0
R:up:34:*:* -10.0
R:up:35:*:* -10.0
R:up:36:*:* -10.0
R:up:37:*:* -10.0
R:up:38:*:* -10.0
R:up:39:*:* -10.0
R:up:40:*:* -10.0
R:up:41:*:* -10.0
R:up:42:*:* -10.0
R:down:0:*:* -10.0
R:down:1:*:* -10.0
R:down:2:*:* -10.0
R:down:3:*:* -10.0
R:down:4:*:* -10.0
R:down:5:*:* -10.0
R:down:6:*:* -10.0
R:down:7:*:* -10.0
R:down:8:*:* -10.0
R:down:9:*:* -10.0
R:down:10:*:* -10.0
R:down:11:*:* -10.0
R:down:12:*:* -10.0
R:down:13:*:* -10.0
R:down:14:*:* -10.0
R:down:15:*:* -10.0
```

```
R:down:16:*:* -10.0
R:down:17:*:* -10.0
R:down:18:*:* -10.0
R:down:19:*:* -10.0
R:down:20:*:* -10.0
R:left:0:*:* -10.0
R:left:1:*:* -10.0
R:left:2:*:* -10.0
R:left:3:*:* -10.0
R:left:4:*:* -10.0
R:left:5:*:* -10.0
R:left:6:*:* -10.0
R:left:21:*:* -10.0
R:left:22:*:* -10.0
R:left:23:*:* -10.0
R:left:24:*:* -10.0
R:left:25:*:* -10.0
R:left:26:*:* -10.0
R:left:27:*:* -10.0
R:right:14:*:* -10.0
R:right:15:*:* -10.0
R:right:16:*:* -10.0
R:right:17:*:* -10.0
R:right:18:*:* -10.0
R:right:19:*:* -10.0
R:right:20:*:* -10.0
R:right:35:*:* -10.0
R:right:36:*:* -10.0
R:right:37:*:* -10.0
R:right:38:*:* -10.0
R:right:39:*:* -10.0
R:right:40:*:* -10.0
R:right:41:*:* -10.0
R:4:*:0:* 5.0
R:4:*:8:* 5.0
R:4:*:16:* 5.0
R:4:*:24:* 5.0
R:4:*:32:* 5.0
R:4:*:40:* 5.0
```

VITA

SHARETHRAM HARIHARAN

Candidate for the Degree of

Master of Science

Thesis: IMPROVING DYNAMIC DECISION-MAKING THROUGH RFID: A PARTIALLY OBSERVABLE MARKOV DECISION PROCESS (POMDP) FOR RFID-ENHANCED WAREHOUSE SEARCH OPERATIONS

Major Field: Industrial Engineering and Management

Biographical:

Personal Data:
Son of Mr. V. Hariharan and Mrs. Sobana Hariharan, born in Chennai, India on July 9^{th}, 1983

Education: Graduated from higher secondary school in Chennai, India; received a bachelor's degree from University of Madras, India in May 2004 in the field of Electronics & Communication Engineering. Completed the requirements for the Master of Science degree with a major in Industrial Engineering and Management at Oklahoma State University in December, 2006.

Experience:
January 2005- December 2006 Graduate Research Assistant in COMMSENS lab at OSU-Stillwater
January 2005- May 2006 Graduate Assistant in CEAT computer labs, OSU

Professional Memberships:
American Society for Quality
Institute for Operations Research and Management Science
Institute of Industrial Engineers
Society of Manufacturing Engineers

Name: Sharethram Hariharan Date of Degree: December2006

Institution: Oklahoma State University Location: Stillwater, Oklahoma

Title of Study: IMPROVING DYNAMIC DECISION-MAKING THROUGH RFID: A PARTIALLY OBSERVABLE MARKOV DECISION PROCESS (POMDP) FOR RFID-ENHANCED WAREHOUSE SEARCH OPERATIONS

Pages in Study: 110 Candidate for the Degree of Master of Science

Major Field: Industrial Engineering and Management

Scope and Method of Study: Misplaced items contribute significantly (2-10%) to the operational expense of a typical warehouse. In this work we develop a Partially Observable Markov Decision Process (POMDP) model for RFID directed search to detect misplaced items within a storage environment. A forklift operator (FLO) equipped with an RFID reader is assigned to search for a misplaced item in a warehouse. A FLO does not know the location of the tagged misplaced item and is guided by the imperfect variations in the strength of the signal received from the RFID tag (active or passive). The model considers five actions, five observations in scenarios with different RFID signal strength distributions namely, excellent, good and poor. An extensive simulation study has been conducted to evaluate the performance of RFID-driven POMDP search method. Specifically, the effects of signal strength distributions, initial beliefs at the start of the search and, the discount factor have been studied.

Findings and Conclusions: The POMDP provides shortest path to locate the tag in the excellent observation scenario. As the observation probabilities decrease, i.e., there is more random imperfections in the nature of signals received (observations), the number of steps to reach the tag increases considerably. The expected reward from a 20-step POMDP with reasonable observation probabilities (varying between 70 and 90%) was 56% higher than that for a no-RFID case. This result implies that a significant (~56%) reduction in search times and efforts for locating a misplaced item are possible using RFID in the tested scenarios. These results, we anticipate, will spur further research on using RFID signals as beacons for searching (perhaps multiple) missing items in more complicated warehouse layouts, through appropriate decomposition algorithms and local, near-optimal policies. Our ongoing research is focuses on achieving these advancements.

ADVISER'S APPROVAL: <u>Dr. Satish Bukkapatnam</u>